

On goodness-of-fit tests for parametric hypotheses in perturbed dynamical systems using a minimum distance estimator

Maroua BEN ABDEDDAIEM

Laboratoire Manceau de Mathématiques, Université du Maine
Le Mans, France

Abstract

We consider the problem of the construction of the Goodness-of-Fit test in the case of continuous time observations of a diffusion process with small noise. The null hypothesis is parametric and we use a minimum distance estimator of the unknown parameter. We propose an asymptotically distribution free test for this model.

Key words: Goodness-of-fit test, minimum distance estimator, asymptotically distribution free tests.

1 Introduction

We consider the problem of the construction of the asymptotically distribution free (ADF) test in the case of continuous time observations of a diffusion process. We suppose that under the null hypothesis the trend coefficient depends on some unknown one-dimensional parameter. Therefore the basic (null) hypothesis is parametric.

The goodness-of-fit (GoF) test plays an important role in mathematical statistics because its application allows to check if the observations correspond well to the proposed mathematical model. Remind that in the case of i.i.d. observations the problem of the construction of GoF tests attracts attention of the statisticians since 1928 due to the works by Cramér (1928), von Mises (1931) and Smirnov (1937) (see, e.g., [2] and [3]).

Let us recall the well-known basic results for the i.i.d. model. Suppose that the null hypothesis is simple. Denoting the continuous distribution function under the null hypothesis by $F_0(x)$, we have to check if the i.i.d.

observations $X^n = (X_1, \dots, X_n)$ have this continuous distribution function. Many GoF tests are based on the following property: the normalized empirical distribution function

$$\sqrt{n} \left(\hat{F}_n(x) - F_0(x) \right), \quad \hat{F}_n(x) = \frac{1}{n} \sum_{j=1}^n \mathbb{I}_{\{X_j < x\}}$$

converges in distribution, under the null hypothesis, to the Brownian bridge $B(F_0(x))$. In particular, for the Cramér-von Mises statistic we have (with the change of variable $s = F_0(x)$)

$$\begin{aligned} \Delta_n &= n \int_{-\infty}^{\infty} \left(\hat{F}_n(x) - F_0(x) \right)^2 dF_0(x) \\ &\implies \int_{-\infty}^{\infty} B(F_0(x))^2 dF_0(x) = \int_0^1 B(s)^2 ds \equiv \Delta. \end{aligned}$$

Therefore the limit Δ does not depend on $F_0(\cdot)$ (distribution free) and the Cramér-von Mises test

$$\psi_n = \mathbb{I}_{\{\Delta_n > c_\alpha\}}, \quad \mathbf{P}\{\Delta > c_\alpha\} = \alpha$$

has the asymptotic ($n \rightarrow \infty$) size $\alpha \in (0, 1)$ (see, e.g., [14]).

In the case of parametric null hypothesis

$$\mathcal{H}_0 : \quad F(x) = F_0(\vartheta, x), \quad \vartheta \in \Theta \subseteq \mathbb{R},$$

where ϑ is some unknown one-dimensional parameter, the similar limit

$$U_n(x) = \sqrt{n} \left(\hat{F}_n(x) - F_0(\hat{\vartheta}_n, x) \right) \implies B(F_0(\vartheta, x)) - \zeta \dot{F}_0(\vartheta, x) \equiv U(x)$$

is no more distribution free and the choice of the threshold c_α is much more complicated. Here $\hat{\vartheta}_n$ is the maximum likelihood estimator (MLE), $\dot{F}_0(\vartheta, \cdot)$ is the derivative of $F_0(\vartheta, \cdot)$ w.r.t. ϑ and ζ is a Gaussian variable. We have for $U(x)$, by the change of variables $t = F_0(\vartheta, x)$ and $s = F_0(\vartheta, y)$, the representation (see, e.g., Darling [2])

$$u(t) = B(t) - \int_0^1 h(\vartheta, s) dB(s) \int_0^t h(\vartheta, s) ds, \quad (1)$$

where

$$h(\vartheta, s) = \frac{\dot{l}(\vartheta, F_\vartheta^{-1}(s))}{\sqrt{I(\vartheta)}}, \quad \int_0^1 h(\vartheta, s)^2 ds = 1.$$

Here $l(\vartheta, x) = \ln f(\vartheta, x)$, $\dot{l}(\vartheta, x)$ is the derivative of $l(\vartheta, x)$ w.r.t. ϑ , $f(\vartheta, x)$ is the density function, $I(\vartheta)$ is the Fisher information and $y = F_\vartheta^{-1}(s)$ is the inverse function to $F_0(\vartheta, y)$, i.e., $F_0(\vartheta, y) = s$. Note that the function $h(\cdot, \cdot)$ appears here twice because we used the MLE.

For the minimum distance estimator (MDE) using the same arguments as in the case of the MLE we obtain the limit

$$u(t) = B(t) - \int_0^1 g(\vartheta, s) dB(s) \int_0^t h(\vartheta, s) ds, \quad \int_0^1 g(\vartheta, s)^2 ds = 1 \quad (2)$$

with two different functions. In both cases, the corresponding tests are no more asymptotically distribution free and the choice of the threshold from the equation

$$\mathbf{P} \left\{ \int_0^1 u(s)^2 ds > c_\alpha \right\} = \alpha$$

can be a difficult problem.

There are several possibilities to solve this problem in the case of the limit (1). One of them is to find a linear transformation $L[\cdot]$ of the random function $u(\cdot)$, such that $L[u](t) = W(t)$, where $W(t)$, $0 \leq t \leq 1$ is some Wiener process. Then the statistic

$$\begin{aligned} \hat{\Delta}_n &= \int_{-\infty}^{\infty} L[U_n](x)^2 dF_0(\hat{\vartheta}_n, x) \implies \int_{-\infty}^{\infty} W(F_0(\vartheta, x))^2 dF_0(\vartheta, x) \\ &= \int_0^1 W(t)^2 dt \equiv \hat{\Delta} \end{aligned}$$

and we obtain once more the distribution free limit $\hat{\Delta}$. Therefore the test

$$\hat{\psi}_n = \mathbb{I}_{\{\hat{\Delta}_n > d_\alpha\}}, \quad \mathbf{P} \left\{ \hat{\Delta} > d_\alpha \right\} = \alpha$$

is ADF. Such transformation was proposed by Khmaladze [6]. We have to emphasize that in [6] and in many other works (see, e.g., the paper Maglapheridze et al. [16]) the estimator used was always the MLE and this is important for the construction of this linear transformation. Many authors wrote that similar transformation can be obtained in the case of other estimators with limit (2), but as we know this work (construction of the linear transformation with other estimators) was not done.

The problem of GoF testing for the model of continuous time observations of diffusion process, with a simple null hypothesis $\Theta = \{\vartheta_0\}$, was studied in [1] and [10]. Suppose that the observed diffusion process under hypothesis is

$$dX_t = S_0(X_t) dt + \varepsilon \sigma(X_t) dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T, \quad (3)$$

with deterministic initial value x_0 , known diffusion coefficient $\varepsilon^2 \sigma(\cdot)^2 > 0$, some known smooth function $S_0(\cdot)$ and $W_t, 0 \leq t \leq 1$ is a Wiener process. Note that if the functions $S_0(x)$ and $\sigma(x)$ are Lipschitz w.r.t. x , then we have with probability 1

$$\sup_{0 \leq t \leq T} |X_t - x_t| \leq C\varepsilon \sup_{0 \leq t \leq T} |W_t|,$$

where C is some constant. Therefore the process X_t converges to $x_t = x_t(\vartheta_0)$ (solution of the equation (3) as $\varepsilon = 0$) uniformly w.r.t. $t \in [0, T]$ with probability 1. We have as well

$$\sup_{0 \leq t \leq T} \mathbf{E} |X_t - x_t|^2 \leq C\varepsilon^2.$$

For the proof see [9], Lemma 1.13.

The GoF test was constructed on the basis of the normalized difference $\varepsilon^{-1}(X_t - x_t)$ and the limit of this statistic is a Gaussian process. This process can be transformed into a Wiener process $w(s), 0 \leq s \leq 1$ as follows.

Introduce the statistic

$$\delta_\varepsilon = \left[\int_0^T \left(\frac{\sigma(x_t)}{S_0(x_t)} \right)^2 dt \right]^{-2} \int_0^T \left(\frac{X_t - x_t}{\varepsilon S_0(x_t)^2} \right)^2 \sigma(x_t)^2 dt.$$

The following convergence

$$\delta_\varepsilon \implies \delta \equiv \int_0^1 w(s)^2 ds$$

was proved and therefore the test $\phi_\varepsilon = \mathbb{I}_{\{\delta_\varepsilon > c_\varepsilon\}}$ with $\mathbf{P}(\delta > c_\varepsilon) = \varepsilon$ is ADF.

The case of parametric basic hypothesis and ADF tests for “small noise” diffusion processes was studied, for example, in [5], [11]–[13] and the estimator used, for the construction of the linear transformation and tests, was always the MLE (see Kutoyants [8]). There are several ADF GoF tests for the ergodic diffusion processes proposed, for example, in the works [17]–[18] and [7].

Note that in some problems it is not possible to have an explicit expression for the MLE and therefore sometimes it is better to use other estimators, which can be easily calculated. For example, this can be the minimum distance estimator, a method of moments estimator or a trajectory fitting estimator and so on. In all such cases, the limit expression for the underlying statistics will be like (2) but with two different functions $h(\cdot, \cdot)$ and $g(\cdot, \cdot)$ (see below).

In this work, we observe continuous time process X^ε , which is the solution of the stochastic differential equation

$$dX_t = S(X_t) dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T,$$

where W_t , $0 \leq t \leq T$ is a Wiener process and $S(x)$ is some unknown smooth function. Based on a minimum distance estimator, we construct an ADF test in the case of parametric null hypothesis

$$\mathcal{H}_0 : \quad S(x) = S(\vartheta, x), \quad \vartheta \in \Theta = (a, b).$$

The main results are presented in Theorem 1 and Theorem 2. In particular, the ADF test is given in Theorem 2 below. We realize the following program. First, we show that the basic statistic

$$u_\varepsilon(t) = \frac{X_t - x_t(\vartheta_\varepsilon^*)}{\varepsilon S(\vartheta_\varepsilon^*, X_t)}, \quad 0 \leq t \leq T$$

(ϑ_ε^* is the MDE) converges to the random process $u(t)$, $0 \leq t \leq T$ (see (14) below). Then we transform $u(t)$ in $U\left(\frac{t}{T}\right)$, $0 \leq t \leq T$ (see (15) below). We obtain for $U(\cdot)$ the following representation

$$U(\nu) = W(\nu) - \int_0^1 g(\vartheta, r) dW(r) \int_0^\nu h(\vartheta, r) dr, \quad \int_0^1 g(\vartheta, r)^2 dr = 1. \quad (4)$$

The last step is to apply the special linear transformation $L[U](\nu) = w_\nu$, $0 \leq \nu \leq 1$, where w_ν , $0 \leq \nu \leq 1$ is a Wiener process (see Theorem 1). This allows us to construct an ADF test (see Theorem 2). The main contribution of this work is the form of this linear transformation.

Then we realize similar transformations with the “empirical” process $u_\varepsilon(\cdot)$, apply the linear transformation $L[\cdot]$ and obtain the convergence

$$\Delta_\varepsilon = \frac{1}{T} \int_0^T L[U_\varepsilon](t)^2 dt \implies \int_0^1 w_\nu^2 d\nu.$$

Here the process $U_\varepsilon(\cdot)$ will be defined in Section 5 by (44). Therefore the test $\psi_\varepsilon = \mathbb{I}_{\{\Delta_\varepsilon > c_\alpha\}}$ will be ADF, because the limit distribution of Δ_ε does not depend on $S(\cdot)$ and ϑ and the test is of asymptotic ($\varepsilon \rightarrow 0$) size $\alpha \in (0, 1)$.

Note that if in our problem we use the MLE of the unknown parameter, then the limit representation is

$$U(\nu) = W(\nu) - \int_0^1 h(\vartheta, s) dW(s) \int_0^\nu h(\vartheta, s) ds, \quad \int_0^1 h(\vartheta, s)^2 ds = 1 \quad (5)$$

and our transformation coincides with the one proposed in [6]. We discuss this case in Section 6.

We have to note that this linear transformation is rather cumbersome and the realization of the test can be a computationally difficult problem too. We suppose that the presented result is of theoretical interest and allows to “close the gap” in this field.

At the same time, we understand that this result is in some sense “negative” and says that if we have no MLE it is better to seek another GoF test, which is ADF. Note as well that in i.i.d. case, even if the estimated parameter is one-dimensional, the reduction of the equation with Brownian bridge (2) to the equation (4), using the relation $(B(t) = W(t) - t W(1))$, leads to the corresponding Fredholm equation (see (27) below) with two-dimensional $g(\cdot)$ and $h(\cdot)$. The expression for the solution of this equation and the form of the linear transformation becomes much more complicated. This is probably the reason why this problem was not considered till now. Our results thus should be understood as a constructive existence result for ADF tests based on the MDE.

2 Minimum distance estimator

Suppose that the continuous time observed process $X^\varepsilon = (X_t, 0 \leq t \leq T)$ is the solution of the stochastic differential equation

$$dX_t = S(X_t) dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T, \quad (6)$$

where W_t , $0 \leq t \leq T$ is a Wiener process, the initial value x_0 is deterministic and the trend coefficient $S(x)$ is some unknown smooth function.

We consider the composite basic hypothesis

$$\mathcal{H}_0 : \quad S(x) = S(\vartheta, x), \quad \vartheta \in \Theta = (a, b),$$

where ϑ is the one-dimensional (unknown) parameter, against alternative $\mathcal{H}_1 : \text{not } \mathcal{H}_0$, i.e., the trend coefficient $S(x)$ in the observed diffusion process (6) does not belong to the parametric family $\{S(\vartheta, x), \vartheta \in \Theta\}$. Therefore the process, under hypothesis \mathcal{H}_0 , has the stochastic differential

$$dX_t = S(\vartheta, X_t) dt + \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T. \quad (7)$$

We are interested in the properties of the test in the asymptotics of *small noise*, i.e., as $\varepsilon \rightarrow 0$. Below and in the sequel the dot means derivation w.r.t. ϑ .

Let us introduce the regularity conditions.

\mathcal{R} . The function $S(\vartheta, x)$ is strictly positive and has two continuous bounded derivatives with respect to ϑ and x .

In the presentation below we suppose that these conditions and the basic hypothesis \mathcal{H}_0 are always fulfilled.

It is known that the solution X_t converges uniformly in $t \in [0, T]$ to the solution $x_t = x_t(\vartheta)$ of the ordinary differential equation

$$\frac{dx_t}{dt} = S(\vartheta, x_t), \quad x_0, \quad 0 \leq t \leq T,$$

where x_0 is the same as in (6) (see [4] or [9], Lemma 1.13 for the proof).

Recall the properties of maximum likelihood and minimum distance estimators of the parameter ϑ . The likelihood ratio function in the case of observations (7) is

$$\ell(\vartheta, X^\varepsilon) = \frac{dP_\vartheta}{dP_0} = \exp \left\{ \int_0^T \frac{S(\vartheta, X_t)}{\varepsilon^2} dX_t - \int_0^T \frac{S(\vartheta, X_t)^2}{2\varepsilon^2} dt \right\}, \quad \vartheta \in \Theta,$$

where P_ϑ and P_0 are the measures induced respectively by the processes (7) and

$$dX_t = \varepsilon dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T,$$

(see [15] for more details).

The MLE $\hat{\vartheta}_\varepsilon$ is solution of the equation

$$\sup_{\vartheta \in \Theta} \ell(\vartheta, X^\varepsilon) = \ell(\hat{\vartheta}_\varepsilon, X^\varepsilon).$$

This estimator is consistent and asymptotically normal (as $\varepsilon \rightarrow 0$)

$$\varepsilon^{-1}(\hat{\vartheta}_\varepsilon - \vartheta) \Longrightarrow \mathcal{N}(0, I(\vartheta)^{-1}), \quad I(\vartheta) = \int_0^T \dot{S}(\vartheta, x_t)^2 dt.$$

Here and in the sequel $x_t = x_t(\vartheta)$. For the proof see, Kutoyants [9].

Introduce the *minimum distance estimator*

$$\vartheta_\varepsilon^* = \arg \min_{\vartheta \in \Theta} \|X - x(\vartheta)\|,$$

where $\|\cdot\|$ is $L^2[0, T]$ norm defined by

$$\|X - x(\vartheta)\|^2 = \int_0^T (X_t - x_t(\vartheta))^2 dt.$$

The properties of the MDE for this model were studied in [9]. The MDE satisfies the *minimum distance equation* (MDEq)

$$\int_0^T (X_t - x_t(\vartheta_\varepsilon^*)) \dot{x}_t(\vartheta_\varepsilon^*) dt = 0.$$

Let us put $u_\varepsilon^* = \varepsilon^{-1} (\vartheta_\varepsilon^* - \vartheta)$. Therefore the MDEq is as follows:

$$\int_0^T (X_t - x_t(\vartheta + \varepsilon u_\varepsilon^*)) \dot{x}_t(\vartheta_\varepsilon^*) dt = 0.$$

Then, by the Taylor formula, we can write

$$\int_0^T \left(X_t - x_t(\vartheta) - \varepsilon u_\varepsilon^* \dot{x}_t(\tilde{\vartheta}) \right) \dot{x}_t(\vartheta_\varepsilon^*) dt = 0,$$

where $|\tilde{\vartheta} - \vartheta| \leq |\vartheta_\varepsilon^* - \vartheta|$ and for u_ε^* we obtain the following representation

$$u_\varepsilon^* = J(\vartheta)^{-1} \int_0^T \varepsilon^{-1} (X_t - x_t(\vartheta)) \dot{x}_t(\vartheta) dt + o(1), \quad (8)$$

where

$$J(\vartheta) = \int_0^T \dot{x}_t(\vartheta)^2 dt. \quad (9)$$

Let us consider the random process $x_t^{(1)}$ defined as the derivative of X_t w.r.t. ε at $\varepsilon = 0$, i.e., we have $\varepsilon^{-1} (X_t - x_t(\vartheta)) \rightarrow x_t^{(1)}$. Here the random process $x_t^{(1)}$ satisfies the linear equation

$$dx_t^{(1)} = S'(\vartheta, x_t) x_t^{(1)} dt + dW_t, \quad x_0^{(1)} = 0, \quad 0 \leq t \leq T,$$

where $S'(\vartheta, x) = \frac{\partial S(\vartheta, x)}{\partial x}$. The process $x_t^{(1)}$ has the representation

$$x_t^{(1)} = S(\vartheta, x_t) \int_0^t \frac{1}{S(\vartheta, x_s)} dW_s \quad (10)$$

(see, e.g., Section 3.3 in [9]). Due to the above representation of $x_t^{(1)}$, the equation (8) becomes

$$\begin{aligned} u_\varepsilon^* &= J(\vartheta)^{-1} \int_0^T x_t^{(1)} \dot{x}_t(\vartheta) dt + o(1) \\ &= J(\vartheta)^{-1} \int_0^T S(\vartheta, x_t) \int_0^t \frac{1}{S(\vartheta, x_s)} dW_s \dot{x}_t(\vartheta) dt + o(1). \end{aligned}$$

Here $o(1)$ means the convergence in probability, i.e., for any $\nu > 0$, we have

$$\lim_{\varepsilon \rightarrow 0} \mathbf{P}_{\vartheta} (|u_{\varepsilon}^* - \rho(\vartheta)| > \nu) = 0, \quad (11)$$

where

$$\rho(\vartheta) = J(\vartheta)^{-1} \int_0^T S(\vartheta, x_t) \int_0^t \frac{1}{S(\vartheta, x_s)} dW_s \dot{x}_t(\vartheta) dt.$$

Therefore $\vartheta_{\varepsilon}^*$ admits the following representation, by Fubini's Theorem,

$$\varepsilon^{-1} (\vartheta_{\varepsilon}^* - \vartheta) = J(\vartheta)^{-1} \int_0^T \frac{1}{S(\vartheta, x_v)} \int_v^T S(\vartheta, x_s) \dot{x}_s(\vartheta) ds dW_v + o(1). \quad (12)$$

Moreover, under conditions of regularity, the estimator $\vartheta_{\varepsilon}^*$ is consistent and asymptotically normal (see Chapter 7 in [9])

$$\mathcal{L}_{\vartheta} \{ \varepsilon^{-1} (\vartheta_{\varepsilon}^* - \vartheta) \} \implies \mathcal{L} \{ \xi \} = \mathcal{N}(0, \sigma^2(\vartheta)),$$

where

$$\sigma^2(\vartheta) = J(\vartheta)^{-2} \int_0^T \frac{1}{S(\vartheta, x_v)^2} \left(\int_v^T S(\vartheta, x_s) \dot{x}_s(\vartheta) ds \right)^2 dv.$$

3 Basic statistic

Our goal is to find the GoF test, which is ADF, i.e., we seek the test statistic whose limit distribution, under hypothesis, does not depend on the underlying model given by the function $S(\vartheta, x)$ and parameter ϑ .

Introduce the statistic

$$\delta_{\varepsilon}^* = \int_0^T \left[\frac{X_t - x_t(\vartheta_{\varepsilon}^*)}{\varepsilon S(\vartheta_{\varepsilon}^*, X_t)} \right]^2 dt.$$

To study it we need the behavior of the difference $X_t - x_t(\vartheta_{\varepsilon}^*)$, which can be described as follows:

$$\begin{aligned} \varepsilon^{-1} (X_t - x_t(\vartheta_{\varepsilon}^*)) &= \varepsilon^{-1} (X_t - x_t(\vartheta)) - \varepsilon^{-1} (\vartheta_{\varepsilon}^* - \vartheta) \dot{x}_t(\vartheta) + o(1) \\ &= x_t^{(1)}(\vartheta) - J(\vartheta)^{-1} \int_0^T x_s^{(1)} \dot{x}_s(\vartheta) ds \dot{x}_t(\vartheta) + o(1) \\ &= S(\vartheta, x_t) \int_0^t \frac{dW_v}{S(\vartheta, x_v)} \\ &\quad - J(\vartheta)^{-1} \int_0^T \frac{1}{S(\vartheta, x_v)} \int_v^T S(\vartheta, x_s) \dot{x}_s(\vartheta) ds dW_v \dot{x}_t(\vartheta) + o(1), \end{aligned}$$

where the process $x_t^{(1)}$ was defined by (10) and the derivative $\dot{x}_t(\vartheta)$ w.r.t. ϑ satisfies the equation

$$\frac{d\dot{x}_t(\vartheta)}{dt} = S'(\vartheta, x_t) \dot{x}_t(\vartheta) + \dot{S}(\vartheta, x_t), \quad \dot{x}_0(\vartheta) = 0.$$

Its solution is the function (it can be found in [11])

$$\dot{x}_t(\vartheta) = S(\vartheta, x_t) \int_0^t \frac{\dot{S}(\vartheta, x_v)}{S(\vartheta, x_v)} dv.$$

Here $o(1)$ is the uniform convergence in probability w.r.t. $t \in [0, T]$, i.e., for any $\nu > 0$, we have

$$\lim_{\varepsilon \rightarrow 0} \mathbf{P}_\vartheta \left(\sup_{t \in [0, T]} |\varepsilon^{-1} (X_t - x_t(\vartheta_\varepsilon^*)) - \mu_t(\vartheta)| > \nu \right) = 0, \quad (13)$$

where

$$\begin{aligned} \mu_t(\vartheta) &= S(\vartheta, x_t) \int_0^t \frac{dW_v}{S(\vartheta, x_v)} \\ &\quad - J(\vartheta)^{-1} \int_0^T \frac{1}{S(\vartheta, x_v)} \int_v^T S(\vartheta, x_s) \dot{x}_s(\vartheta) ds dW_v \dot{x}_t(\vartheta). \end{aligned}$$

For the details see [9], Chapter 7.

Hence we have the uniform convergence w.r.t. $t \in [0, T]$ (in probability)

$$\begin{aligned} u_\varepsilon(t) &= \frac{X_t - x_t(\vartheta_\varepsilon^*)}{\varepsilon S(\vartheta_\varepsilon^*, X_t)} \longrightarrow u(t) = \int_0^t \frac{dW_v}{S(\vartheta, x_v)} \\ &\quad - J(\vartheta)^{-1} \int_0^T \frac{1}{S(\vartheta, x_v)} \int_v^T S(\vartheta, x_s) \dot{x}_s(\vartheta) ds dW_v \int_0^t \frac{\dot{S}(\vartheta, x_v)}{S(\vartheta, x_v)} dv \end{aligned} \quad (14)$$

and it can be shown (see in [11] the details in the similar problem where the MLE was used) that

$$\delta_\varepsilon^* \Longrightarrow \int_0^T u(t)^2 dt.$$

Therefore the test based on this statistic is not ADF. Hence to obtain an ADF GoF test we introduce the Gaussian process

$$U\left(\frac{t}{T}\right) = \frac{1}{\sqrt{T}} \int_0^t S(\vartheta, x_s) du(s), \quad 0 \leq t \leq T, \quad (15)$$

where $u(\cdot)$ was given by (14). Then, by Itô formula,

$$d(S(\vartheta, x_s)u(s)) = dS(\vartheta, x_s) u(s) + du(s) S(\vartheta, x_s)$$

and using the equality

$$dS(\vartheta, x_s) = S'(\vartheta, x_s) S(\vartheta, x_s) ds, \quad (16)$$

we have

$$S(\vartheta, x_t) u(t) = \int_0^t S'(\vartheta, x_s) S(\vartheta, x_s) u(s) ds + \int_0^t S(\vartheta, x_s) du(s).$$

Here and in the sequel we denoted by prime the derivative w.r.t. x .

Therefore the process $U(\cdot)$ defined by (15) admits the following representation ($0 \leq t \leq T$)

$$U\left(\frac{t}{T}\right) = \frac{1}{\sqrt{T}} S(\vartheta, x_t) u(t) - \frac{1}{\sqrt{T}} \int_0^t S'(\vartheta, x_s) S(\vartheta, x_s) u(s) ds. \quad (17)$$

Further we define two functions

$$h(\vartheta, r) = T \tilde{J}(\vartheta)^{-1} \dot{S}(\vartheta, x_{rT}(\vartheta)) C(\vartheta)^{1/2} \quad (18)$$

and

$$g(\vartheta, r) = S(\vartheta, x_{rT}(\vartheta))^{-1} \int_r^1 S(\vartheta, x_{zT}(\vartheta)) \dot{x}_{zT}(\vartheta) dz C(\vartheta)^{-1/2}. \quad (19)$$

Here $\tilde{J}(\vartheta) = \int_0^1 \dot{x}_{vT}(\vartheta)^2 dv$ and

$$C(\vartheta) = \int_0^1 S(\vartheta, x_{vT}(\vartheta))^{-2} \left(\int_v^1 S(\vartheta, x_{zT}(\vartheta)) \dot{x}_{zT}(\vartheta) dz \right)^2 dv.$$

Observe that

$$\int_0^1 g(\vartheta, r)^2 dr = 1.$$

Lemma 1 *We have the equality*

$$U(\nu) = W(\nu) - \int_0^1 g(\vartheta, r) dW(r) \int_0^\nu h(\vartheta, r) dr, \quad 0 \leq \nu \leq 1, \quad (20)$$

where $W(\nu), 0 \leq \nu \leq 1$ is a Wiener process.

Proof. We have by Itô formula, for $0 \leq t \leq T$,

$$\begin{aligned} dU\left(\frac{t}{T}\right) &= \frac{1}{\sqrt{T}} dS(\vartheta, x_t) u(t) \\ &\quad + \frac{1}{\sqrt{T}} S(\vartheta, x_t) du(t) - \frac{1}{\sqrt{T}} S'(\vartheta, x_t) S(\vartheta, x_t) u(t) dt. \end{aligned}$$

Then by (16) and using the representation (14) of the process $u(\cdot)$, we obtain

$$\begin{aligned} U\left(\frac{t}{T}\right) &= \frac{1}{\sqrt{T}} \int_0^t S(\vartheta, x_s) du(s) \\ &= \frac{W_t}{\sqrt{T}} - \frac{J(\vartheta)^{-1}}{\sqrt{T}} \int_0^T \frac{1}{S(\vartheta, x_v)} \int_v^T S(\vartheta, x_s) \dot{x}_s(\vartheta) ds dW_v \int_0^t \dot{S}(\vartheta, x_v) dv. \end{aligned}$$

Hence, by the change of variables $\nu = \frac{t}{T}$ and $W(\nu) = T^{-1/2}W_{\nu T}$, $0 \leq \nu \leq 1$, we have

$$\begin{aligned} U(\nu) &= W(\nu) \\ &\quad - \frac{J(\vartheta)^{-1}}{\sqrt{T}} \int_0^T \frac{1}{S(\vartheta, x_v)} \int_v^T S(\vartheta, x_s) \dot{x}_s(\vartheta) ds dW_v \int_0^{\nu T} \dot{S}(\vartheta, x_v) dv. \end{aligned}$$

Let us change the variables $r = \frac{v}{T}$, $z = \frac{s}{T}$, $W(r) = T^{-1/2}W_{rT}$, $0 \leq r \leq 1$. Then we can write

$$\begin{aligned} U(\nu) &= W(\nu) \\ &\quad - \tilde{J}(\vartheta)^{-1} \int_0^1 \frac{1}{S(\vartheta, x_{rT})} \int_{rT}^T S(\vartheta, x_s) \dot{x}_s(\vartheta) ds dW(r) \int_0^\nu \dot{S}(\vartheta, x_{rT}) dr \\ &= W(\nu) \\ &\quad - \frac{T}{\tilde{J}(\vartheta)} \int_0^1 \frac{1}{S(\vartheta, x_{rT})} \int_r^1 S(\vartheta, x_{zT}) \dot{x}_{zT}(\vartheta) dz dW(r) \int_0^\nu \dot{S}(\vartheta, x_{rT}) dr \\ &= W(\nu) - \int_0^1 g(\vartheta, r) dW(r) \int_0^\nu h(\vartheta, r) dr. \end{aligned}$$

Therefore we obtain the representation (20) and this proves the Lemma 1.

It can be shown using the convergence of the *empirical version* $U_\varepsilon(\cdot)$ to $U(\cdot)$ (see proof of Theorem 2 below) and due to the continuous mapping Theorem, that we have the convergence

$$\tilde{\Delta}_\varepsilon = \frac{1}{T^2} \int_0^T \left(\int_0^t S(\vartheta_\varepsilon^*, X_s) du_\varepsilon(s) \right)^2 dt \implies \int_0^1 U(\nu)^2 d\nu.$$

We remark that the test based on this statistic is not ADF. Hence we have to find the transformation $L[U](\cdot)$ into the Wiener process such that

$$\int_0^1 L[U](\nu)^2 d\nu = \int_0^1 w_\nu^2 d\nu, \quad 0 \leq \nu \leq 1,$$

where $w_\nu, 0 \leq \nu \leq 1$ is a Wiener process. This property allows us to present the “empirical version” of the test statistic with the same limit. Therefore the test based on this statistic is ADF.

4 Linear transformation

Now the problem is to find such transformation $L[\cdot]$ of $U(\cdot)$ (see (20)) using the MDE that $L[U](\nu) = w_\nu$. Recall that for the limit process (1) such linear transformation and the corresponding ADF test were proposed by Khmaladze [6]. Another (direct) proof of this result was recently obtained by Kleptsyna and Kutoyants [7]. Note that in these works the estimator used was always the MLE and in our work it is the MDE. The limit processes in these two cases are quite different. That is why we have to present here a special modification of the proof given in [7]. Our proof follows the main steps of the work [7]. Specifically, we have to solve Fredholm equation of the second kind with degenerated kernel. The solution of it gives us the desired linear transformation.

Denote

$$\begin{aligned} I_1 &= \int_0^r g(\vartheta, q)^2 dq, \quad I_2 = \int_0^r h(\vartheta, q) g(\vartheta, q) dq, \quad I_3 = \int_0^r h(\vartheta, q) dq, \\ I_4 &= \int_0^r h(\vartheta, q)^2 dq, \quad I_5 = \int_0^r g(\vartheta, q) dq. \end{aligned} \quad (21)$$

Below we omit ϑ and r for simplicity and put $g = g(\vartheta, r)$ and $h = h(\vartheta, r)$. Introduce the functions

$$\begin{aligned} \varphi_1(r) &= g - h - 3I_2g + I_5hg + I_3g^2 + 2I_2h - 2I_2I_3g^2 + I_1I_2^2h + I_4I_5g^2 - I_2^3g \\ &\quad - I_2I_4g + 3I_2^2g + I_2I_5h^2 - 2I_2I_5hg - 2I_1I_2h + I_2^2I_5hg + I_1^2I_3h^2 - I_4h \\ &\quad + 2I_1I_4h - I_1I_4g + I_1I_2I_4g + I_1I_4I_5hg - I_1^2I_4h + I_1h + 2I_2I_3hg \\ &\quad - I_2I_4I_5g^2 - I_5h^2 + 2I_1I_3hg - 2I_1I_2I_3hg - 2I_1I_3h^2 - I_2^2h + I_3h^2 \\ &\quad - 2I_3hg - I_4I_5hg - I_1I_2I_5h^2 + I_2^2I_3g^2 + I_1I_5h^2 + I_4g, \end{aligned} \quad (22)$$

$$\varphi_2(r) = 1 + I_5h - 3I_2I_5h + I_1I_3h + I_3g - 3I_2I_3g + I_4I_5g - I_3h - I_1I_4^2I_5g$$

$$\begin{aligned}
& + 3I_2^2 I_3 g - 2I_2 I_4 I_5 g + 2I_2 I_3 h - I_2^3 I_5 h + I_1 I_2^2 I_3 h - I_2^3 I_3 g + 3I_2^2 I_5 h \\
& + I_4 I_5 h - I_2 I_4 I_5 h + 2I_1 I_3 I_4 h + I_3 I_4 g - I_2 I_3 I_4 g + I_4^2 + 2I_2^2 I_4 + I_2^4 \\
& + I_1 I_2 I_4 I_5 h - I_1^2 I_3 I_4 h - I_1 I_3 I_4 g + I_1 I_2 I_3 I_4 g - I_2^2 I_3 h - 2I_1 I_2^2 I_4 \\
& - 2I_1 I_4^2 + I_1^2 I_4^2 + 2I_4 - 2I_1 I_4 - 4I_2 I_4 + 4I_1 I_2 I_4 - 4I_2 + I_4^2 I_5 g \\
& + I_2^2 I_4 I_5 g - 2I_1 I_2 I_3 h - I_3 I_4 h + 6I_2^2 - 4I_2^3 - I_1 I_4 I_5 h
\end{aligned} \tag{23}$$

and

$$\begin{aligned}
\psi_2(r) = & h + I_3 h g - 2I_2 I_4 g + I_5 h^2 - 3I_2 h - 2I_2 I_3 h g - I_1 I_2 I_3 h^2 + I_4 g + 3I_2^2 h \\
& + I_2^2 I_4 g + I_2^2 I_5 h^2 - I_2^3 h - I_3 I_4 h g + I_4^2 g + I_4 h - I_2 I_4 h - 2I_2 I_5 h^2 \\
& - I_1 I_4 h + I_1 I_2 I_4 h + I_1 I_3 h^2 + I_3 I_4 g^2 + I_2 I_3 h^2 - I_3 h^2 + I_1 I_3 I_4 h g \\
& - I_2 I_3 I_4 g^2 - 2I_2 I_4 I_5 h g + 2I_4 I_5 h g + I_2^2 I_3 h g - I_1 I_4^2 g + I_4^2 I_5 g^2.
\end{aligned} \tag{24}$$

The following Theorem is the main result of this work.

Theorem 1 Suppose that $h(q)$ and $g(q)$ are continuous functions such that $\int_0^1 g(q)^2 dq = 1$ and $\varphi_2(r)$ is a strictly positive function on $[0, 1)$. Then the equality

$$L[U](\nu) = U(\nu) + \int_0^\nu \int_0^r \frac{\varphi_1(r) h(q) + \psi_2(r) g(q)}{\varphi_2(r)} dU(q) dr = w_\nu \tag{25}$$

holds. Here w_ν , $0 \leq \nu \leq 1$ is a Wiener process.

Proof. The proof will be done in several steps.

Step 1: Introduce a Gaussian process

$$M_t = \int_0^t q(t, s) dU(s), \quad 0 \leq t \leq 1,$$

where the function $q(t, s)$ is chosen as solution of Fredholm equation described in the next step. Observe that

$$M_t = \int_0^t q(t, u) dW(u) - \int_0^1 g(u) dW(u) \int_0^t q(t, u) h(u) du.$$

Step 2: For the correlation function of M_t

$$R(t, s) = \mathbf{E}[M_t M_s], \quad t > s,$$

we have

$$\mathbf{E}[M_t M_s] = \mathbf{E} \left[\int_0^t q(t, u) dW(u) - \int_0^1 g(u) dW(u) \int_0^t q(t, u) h(u) du \right]$$

$$\begin{aligned}
& \left[\int_0^s q(s, v) \, dW(v) - \int_0^1 g(v) \, dW(v) \int_0^s q(s, v) h(v) \, dv \right] \\
&= \int_0^s q(t, u) q(s, u) \, du - \int_0^t q(t, u) g(u) \, du \int_0^s q(s, v) h(v) \, dv \\
&\quad - \int_0^s q(s, v) g(v) \, dv \int_0^t q(t, u) h(u) \, du \\
&\quad + \underbrace{\int_0^1 g(v)^2 \, dv}_{=1} \int_0^t q(t, u) h(u) \, du \int_0^s q(s, v) h(v) \, dv \\
&= \int_0^s q(s, u) \left[q(t, u) - \int_0^t q(t, v) g(v) \, dv h(u) \right. \\
&\quad \left. - \int_0^t q(t, v) h(v) \, dv g(u) + \int_0^t q(t, v) h(v) \, dv h(u) \right] du \\
&= \int_0^s q(s, u) \left[q(t, u) - \int_0^t q(t, v) [g(v) h(u) \right. \\
&\quad \left. + h(v) g(u) - h(v) h(u)] \, dv \right] du.
\end{aligned}$$

Denote the kernel

$$K(u, v) = g(v) h(u) + h(v) g(u) - h(v) h(u).$$

Then

$$\mathbf{E}[M_t M_s] = \int_0^s q(s, u) \left[q(t, u) - \int_0^t q(t, v) K(u, v) \, dv \right] du. \quad (26)$$

Therefore if we take $q(t, s)$ such that it solves the Fredholm equation of the second kind (t is fixed)

$$q(t, s) - \int_0^t q(t, v) K(s, v) \, dv = 1, \quad s \in [0, t], \quad (27)$$

then (26) becomes

$$\mathbf{E}[M_t M_s] = \mathbf{E}[M_s^2] = \int_0^s q(s, u) \, du. \quad (28)$$

Step 3: The solution $q(t, s)$ can be found as follows. We have

$$q(t, s) = 1 + \int_0^t q(t, v) K(s, v) \, dv.$$

Denote

$$A(t) = \int_0^t q(t, v) h(v) dv$$

and

$$B(t) = \int_0^t q(t, v) g(v) dv.$$

Then $q(t, s)$ has the representation

$$q(t, s) = 1 + B(t) h(s) + A(t) (g(s) - h(s)), \quad (29)$$

where the function $A(t)$ itself is solution of the following equation (after multiplying (29) by $h(s)$ and integrating)

$$\begin{aligned} \int_0^t h(s) ds &= A(t) - B(t) \int_0^t h(s)^2 ds \\ &\quad - A(t) \int_0^t h(s) g(s) ds + A(t) \int_0^t h(s)^2 ds. \end{aligned} \quad (30)$$

The function $B(t)$ is solution of the following equation (after multiplying (29) by $g(s)$ and integrating)

$$\begin{aligned} \int_0^t g(s) ds &= B(t) - B(t) \int_0^t h(s) g(s) ds \\ &\quad - A(t) \int_0^t g(s)^2 ds + A(t) \int_0^t h(s) g(s) ds. \end{aligned} \quad (31)$$

Using the notation (21), we can write (30)–(31) as follows:

$$A(t) - B(t) I_4 - A(t) I_2 + A(t) I_4 = I_3 \quad (32)$$

and

$$B(t) - B(t) I_2 - A(t) I_1 + A(t) I_2 = I_5. \quad (33)$$

Further, we have to find the expressions of $A(t)$ and $B(t)$. Therefore we obtain from (33)

$$B(t) = \frac{I_5 + A(t) (I_1 - I_2)}{(1 - I_2)}. \quad (34)$$

Then we insert (34) in (32) and obtain

$$A(t) = \frac{I_3 (1 - I_2) + I_4 I_5}{(1 - I_2)^2 + I_4 I_6}$$

and

$$B(t) = \frac{I_5 (1 + I_4 - I_2) + I_3 (I_1 - I_2)}{(1 - I_2)^2 + I_4 I_6},$$

where

$$I_6 = \int_t^1 g(s)^2 ds.$$

Therefore the solution $q(t, s)$ of (27) is

$$\begin{aligned} q(t, s) = 1 + & \frac{I_5 (1 + I_4 - I_2) + I_3 (I_1 - I_2)}{(1 - I_2)^2 + I_4 I_6} h(s) \\ & + \frac{I_3 (1 - I_2) + I_4 I_5}{(1 - I_2)^2 + I_4 I_6} (g(s) - h(s)). \end{aligned}$$

The final expression of $q(t, s)$ is

$$\begin{aligned} q(t, s) = 1 + & \frac{\int_0^t g(s) ds \left(1 + \int_0^t h(s)^2 ds - \int_0^t h(s)g(s) ds \right) h(s)}{\left(1 - \int_0^t h(s)g(s) ds \right)^2 + \int_0^t h(s)^2 ds \int_t^1 g(s)^2 ds} \\ & + \frac{\int_0^t h(s) ds \left(\int_0^t g(s)^2 ds - \int_0^t h(s)g(s) ds \right) h(s)}{\left(1 - \int_0^t h(s)g(s) ds \right)^2 + \int_0^t h(s)^2 ds \int_t^1 g(s)^2 ds} \\ & + \frac{\int_0^t h(s) ds \left(1 - \int_0^t h(s)g(s) ds \right) + \int_0^t h(s)^2 ds \int_0^t g(s) ds}{\left(1 - \int_0^t h(s)g(s) ds \right)^2 + \int_0^t h(s)^2 ds \int_t^1 g(s)^2 ds} (g(s) - h(s)) \end{aligned}$$

Step 4: To show that M_t is martingale we need the following Lemma.

Lemma 2 *We have the following equality*

$$\int_0^t q(t, s) ds = \int_0^t q(s, s)^2 ds, \quad 0 \leq t \leq 1. \quad (35)$$

Proof. Show that

$$\frac{d}{dt} \int_0^t q(t, s) ds = \frac{d}{dt} \int_0^t q(s, s)^2 ds = q(t, t)^2.$$

Denote

$$\begin{aligned}
C(t) &= 1 - \int_0^t h(s) g(s) \, ds, \\
D(t) &= \int_0^t g(s) \, ds \left(1 + \int_0^t h(s)^2 \, ds - \int_0^t h(s) g(s) \, ds \right) \\
&\quad + \int_0^t h(s) \, ds \left(\int_0^t g(s)^2 \, ds - \int_0^t h(s) g(s) \, ds \right), \\
K(t) &= C(t)^2 + \int_0^t h(s)^2 \, ds \int_t^1 g(s)^2 \, ds
\end{aligned}$$

and

$$N(t) = C(t) \int_0^t h(s) \, ds + \int_0^t h(s)^2 \, ds \int_0^t g(s) \, ds.$$

Then $q(t, s)$ has the following expression

$$q(t, s) = 1 + \frac{D(t)}{K(t)} h(s) + \frac{N(t)}{K(t)} (g(s) - h(s)).$$

Hence

$$\begin{aligned}
\frac{d}{dt} \int_0^t q(t, s) \, ds &= 1 + \frac{D(t) h(t) + N(t) (g(t) - h(t))}{K(t)} \\
&\quad + \frac{(D'(t) K(t) - D(t) K'(t))}{K(t)^2} \int_0^t h(s) \, ds \\
&\quad + \frac{(N'(t) K(t) - N(t) K'(t))}{K(t)^2} \int_0^t (g(s) - h(s)) \, ds.
\end{aligned} \tag{36}$$

Then we obtain the equalities

$$\begin{aligned}
C(t) &= 1 - I_2, \\
D(t) &= I_5 (1 + I_4 - I_2) + I_3 (I_1 - I_2), \\
K(t) &= (1 - I_2)^2 + I_4 (1 - I_1), \\
N(t) &= I_3 (1 - I_2) + I_4 I_5.
\end{aligned}$$

The derivatives of these functions w.r.t. t have such expressions

$$\begin{aligned}
C'(t) &= -h(t) g(t), \\
D'(t) &= g(t) + I_4 g(t) - I_2 g(t) + I_5 h(t)^2 - I_5 h(t) g(t) \\
&\quad + I_1 h(t) - I_2 h(t) + I_3 g(t)^2 - I_3 h(t) g(t),
\end{aligned}$$

$$\begin{aligned}
K'(t) &= -2h(t) g(t) + 2I_2 h(t)g(t) + h(t)^2 - I_1 h(t)^2 - I_4 g(t)^2, \\
N'(t) &= -I_3 h(t) g(t) + I_4 g(t) + I_5 h(t)^2 + h(t) - I_2 h(t).
\end{aligned}$$

Thus, (36) has the following representation

$$\frac{d}{dt} \int_0^t q(t, s) ds = 1 + \frac{\Phi_1(t) + \Phi_2(t) + \Phi_3(t)}{K(t)^2}, \quad (37)$$

where

$$\begin{aligned}
\Phi_1(t) &= (D(t) h(t) + N(t) (g(t) - h(t))) K(t), \\
\Phi_2(t) &= (D'(t) K(t) - D(t) K'(t)) I_3
\end{aligned}$$

and

$$\Phi_3(t) = (N'(t) K(t) - N(t) K'(t)) (I_5 - I_3).$$

Then returning to initial notation we obtain the expressions

$$\begin{aligned}
\Phi_1(t) &= 3I_2^2 I_5 h - 3I_2 I_5 h + I_1 I_3 h - 3I_2 I_3 g + I_1 I_2^2 I_3 h - I_2^3 I_5 h + I_3 g - I_2^3 I_3 g \\
&\quad + I_4 I_5 g - I_1 I_3 I_4 g - I_3 h - I_1^2 I_3 I_4 h + 2I_2 I_3 h + I_5 h - I_1 I_4^2 I_5 g - I_2 I_3 I_4 g \\
&\quad - 2I_1 I_2 I_3 h + 3I_2^2 I_3 g - 2I_2 I_4 I_5 g + I_4 I_5 h - I_2 I_4 I_5 h + I_1 I_2 I_3 I_4 g + I_4^2 I_5 g \\
&\quad + I_2^2 I_4 I_5 g - I_2^2 I_3 h + 2I_1 I_3 I_4 h - I_3 I_4 h - I_1 I_4 I_5 h + I_3 I_4 g + I_1 I_2 I_4 I_5 h, \\
\Phi_2(t) &= I_2 I_3^2 h^2 + 2I_3 I_4 g - 3I_2 I_3 g + I_3 I_5 h g + I_3 I_4^2 I_5 g^2 + I_3^2 I_4 g^2 + I_1 I_3 h \\
&\quad - 3I_2 I_3 I_4 g + 3I_2^2 I_3 g - I_2 I_3 I_5 h^2 - 2I_2 I_3 I_5 h g + I_3^2 g^2 - I_1 I_3^2 h^2 \\
&\quad + I_2^2 I_3 I_4 g - I_2^2 I_3 g + I_2^2 I_3 I_5 h^2 + I_2^2 I_3 I_5 h g + I_1 I_2^2 I_3 h - 2I_1 I_2 I_3 h \\
&\quad - I_3^2 h g - 2I_2 I_3^2 g^2 + I_2^2 I_3^2 g^2 + 2I_2^2 I_3 h - 2I_2 I_3 I_4 I_5 h g - I_1^2 I_3 I_4 h \\
&\quad + I_2^2 I_3^2 h g + I_3 I_4^2 g + I_3 I_4 I_5 h g - 2I_1 I_2 I_3^2 h g + I_1^2 I_3^2 h^2 - I_1 I_2 I_3^2 h^2 \\
&\quad - I_1 I_3 I_4 g - I_1 I_3 I_4^2 g + I_1 I_2 I_3 I_4 g + I_1 I_3 I_4 I_5 h g - I_1 I_2 I_3 I_5 h^2 + I_3 g \\
&\quad + I_1 I_3^2 I_4 h g + I_1 I_3 I_5 h^2 + I_3 I_4 I_5 g^2 + I_1 I_3 I_4 h - I_2 I_3 h - I_2 I_3^2 I_4 g^2 \\
&\quad - I_2 I_3 I_4 I_5 g^2 + 2I_1 I_3^2 h g - I_3^2 I_4 h g + I_1 I_2 I_3 I_4 h - I_2 I_3 I_4 h - I_2^3 I_3 h
\end{aligned}$$

and

$$\begin{aligned}
\Phi_3(t) &= I_3 I_5 h g + I_2 I_3 I_4 h - I_3 I_5 h^2 - 3I_2 I_5 h - I_3 I_4^2 g - I_2^2 I_3^2 h g \\
&\quad - 2I_2 I_5^2 h^2 + 3I_2^2 I_5 h + I_2^2 I_3 I_5 h g - I_2^2 I_3 I_5 h^2 + I_2^2 I_5^2 h^2 - I_3 I_4^2 I_5 g^2 \\
&\quad + I_4^2 I_5 g - 3I_2^2 I_3 h - I_2 I_4 I_5 h + I_1 I_3 I_4 I_5 h g - I_3^2 h g - I_1 I_4^2 I_5 g \\
&\quad + I_2 I_3^2 I_4 g^2 + I_1 I_3 I_5 h^2 + I_3 I_4 I_5 g^2 + I_2 I_3 I_5 h^2 + I_5 h - 2I_2 I_3 I_5 h g \\
&\quad + 2I_4 I_5^2 h g - 2I_2 I_4 I_5^2 h g + I_4^2 I_5^2 g^2 + I_2^2 I_4 I_5 g - I_3 h + I_1 I_2 I_4 I_5 h \\
&\quad + 3I_2 I_3 h + 2I_2 I_3^2 h g + 2I_2 I_3 I_4 g + 2I_2 I_3 I_5 h^2 + I_5^2 h^2 - I_1 I_4 I_5 h
\end{aligned}$$

$$\begin{aligned}
& -I_2^2 I_3 I_4 g + I_2^3 I_3 h + I_3^2 I_4 h g + I_4 I_5 h - I_2^3 I_5 h + I_3^2 h^2 + I_4 I_5 g \\
& -I_1 I_3^2 I_4 h g + I_1 I_2 I_3^2 h^2 + I_1 I_3 I_4^2 g + I_1 I_3 I_4 h - I_3 I_4 h - I_3 I_4 g \\
& -I_3^2 I_4 g^2 - I_2 I_3^2 h^2 - I_3 I_5 h^2 - I_1 I_2 I_3 I_4 h - I_2 I_3 I_4 I_5 g^2 - I_1 I_3^2 h^2 \\
& -2 I_3 I_4 I_5 h g + 2 I_2 I_3 I_4 I_5 h g - 2 I_2 I_4 I_5 g - I_1 I_2 I_3 I_5 h^2 - I_3 I_4 I_5 h g.
\end{aligned}$$

Denote

$$\Phi(t) = \Phi_1(t) + \Phi_2(t) + \Phi_3(t),$$

then we can write

$$\begin{aligned}
\Phi(t) = & h(t) \left(2I_5 - 6I_2 I_5 + 2I_1 I_3 - 2I_3 + 6I_2^2 I_5 - 4I_1 I_2 I_3 + 4I_2 I_3 + 4I_1 I_3 I_4 \right. \\
& - 2I_2^3 I_5 + 2I_1 I_2^2 I_3 + 2I_4 I_5 - 2I_2 I_4 I_5 - 2I_3 I_4 + 2I_1 I_2 I_4 I_5 - 2I_1^2 I_3 I_4 \\
& \left. - 2I_1 I_4 I_5 - 2I_2^2 I_3 \right) + g(t) \left(2I_3 - 2I_1 I_3 I_4 + 2I_1 I_2 I_3 I_4 - 6I_2 I_3 + 2I_4 I_5 \right. \\
& - 2I_1 I_4^2 I_5 + 6I_2^2 I_3 - 4I_2 I_4 I_5 - 2I_2^3 I_3 + 2I_2^2 I_4 I_5 + 2I_3 I_4 + 2I_4^2 I_5 \\
& \left. - 2I_2 I_3 I_4 \right) + h(t)g(t) \left(2I_3 I_5 - 2I_3^2 - 4I_2 I_3 I_5 - 2I_3 I_4 I_5 + 2I_1 I_3 I_4 I_5 \right. \\
& + 2I_2 I_3^2 + 2I_4 I_5^2 + 2I_1 I_3^2 - 2I_1 I_2 I_3^2 - 2I_2 I_4 I_5^2 + 2I_2^2 I_3 I_5 \left. \right) + h(t)^2 \\
& \left(2I_1 I_3 I_5 - 2I_1 I_2 I_3 I_5 - 2I_2 I_5^2 + I_2^2 I_5^2 - 2I_3 I_5 + 2I_2 I_3 I_5 - 2I_1 I_3^2 + I_3^2 \right. \\
& \left. + I_5^2 + I_1^2 I_3^2 \right) + g(t)^2 \left(I_3^2 - 2I_2 I_3^2 + 2I_3 I_4 I_5 - 2I_2 I_3 I_4 I_5 + I_4^2 I_5^2 + I_2^2 I_3^2 \right).
\end{aligned}$$

Finally, (37) can be written as follows:

$$\frac{d}{dt} \int_0^t q(t, s) \, ds = 1 + \frac{\Phi(t)}{K(t)^2}. \quad (38)$$

Now the expression of $q(t, t)^2$ is

$$\begin{aligned}
q(t, t)^2 &= \left(1 + \frac{D(t)}{K(t)} h(t) + \frac{N(t)}{K(t)} (g(t) - h(t)) \right)^2 \\
&= 1 + \frac{2 h(t) K(t) (D(t) - N(t)) + 2 g(t) K(t) N(t)}{K(t)^2} \\
&\quad + \frac{h(t)^2 (D(t) - N(t))^2 + g(t)^2 N(t)^2}{K(t)^2} \\
&\quad + \frac{2 h(t) g(t) N(t) (D(t) - N(t))}{K(t)^2}.
\end{aligned}$$

Denote

$$M(t) = 2 K(t) (D(t) - N(t)),$$

$$\begin{aligned}
Q(t) &= 2 K(t) N(t), \\
L(t) &= (D(t) - N(t))^2, \\
E(t) &= N(t)^2, \\
H(t) &= 2 N(t)(D(t) - N(t)).
\end{aligned}$$

Therefore the final expression for $q(t, t)^2$ is

$$\begin{aligned}
q(t, t)^2 &= 1 + \frac{h(t) M(t) + g(t) Q(t) + h(t)^2 L(t)}{K(t)^2} \\
&\quad + \frac{g(t)^2 E(t) + h(t) g(t) H(t)}{K(t)^2},
\end{aligned} \tag{39}$$

where

$$\begin{aligned}
M(t) &= 2I_1 I_3 - 6I_2 I_5 - 2I_3 + 6I_2^2 I_5 - 4I_1 I_2 I_3 + 4I_2 I_3 - 2I_2^3 I_5 + 2I_4 I_5 \\
&\quad + 2I_1 I_2^2 I_3 - 2I_1^2 I_3 I_4 - 2I_2^2 I_3 - 2I_2 I_4 I_5 + 2I_5 - 2I_3 I_4 - 2I_1 I_4 I_5 \\
&\quad + 2I_1 I_2 I_4 I_5 + 4I_1 I_3 I_4, \\
Q(t) &= 2I_3 I_4 - 6I_2 I_3 + 6I_2^2 I_3 - 2I_2^3 I_3 + 2I_4 I_5 - 4I_2 I_4 I_5 + 2I_2^2 I_4 I_5 \\
&\quad + 2I_3 - 2I_2 I_3 I_4 - 2I_1 I_3 I_4 + 2I_1 I_2 I_3 I_4 + 2I_4^2 I_5 - 2I_1 I_4^2 I_5, \\
L(t) &= I_5^2 + I_2^2 I_5^2 - 2I_2 I_5^2 + I_1^2 I_3^2 + I_3^2 - 2I_1 I_3^2 + 2I_1 I_3 I_5 - 2I_3 I_5 \\
&\quad + 2I_2 I_3 I_5 - 2I_1 I_2 I_3 I_5, \\
E(t) &= I_3^2 - 2I_2 I_3^2 + I_2^2 I_3^2 + I_4^2 I_5^2 + 2I_3 I_4 I_5 - 2I_2 I_3 I_4 I_5, \\
H(t) &= 2I_3 I_5 - 4I_2 I_3 I_5 - 2I_3 I_4 I_5 + 2I_1 I_3^2 + 2I_2 I_3^2 + 2I_2^2 I_3 I_5 \\
&\quad - 2I_2 I_4 I_5^2 + 2I_1 I_3 I_4 I_5 - 2I_3^2 - 2I_1 I_2 I_3^2 + 2I_4 I_5^2.
\end{aligned}$$

The comparison of these expressions with (38)–(39) shows that the Lemma is proved.

Step 5: In the next step, we need the following Lemma to show that the linear transformation is a Wiener process.

Lemma 3 *If the Gaussian process M_s satisfies (28) and we have relation (35), then*

$$z(t) = \int_0^t q(s, s)^{-1} dM_s$$

is a Wiener process.

Proof. The proof can be found, e.g., in [7], Lemma 2.

Hence

$$M_t = \int_0^t q(s, s) dw_s, \quad 0 \leq t \leq 1$$

is a Gaussian martingale, where $w_s, 0 \leq s \leq 1$ is a Wiener process.

Therefore we have the equality

$$w_t = \int_0^t q(s, s)^{-1} dM_s = U(t) + \int_0^t q(s, s)^{-1} \int_0^s q'_s(s, v) dU(v) ds,$$

where $w_t, 0 \leq t \leq 1$ is a Wiener process (by Lemma 3).

Now we have to calculate the right side of the above expression. The derivative $q'_t(t, s)$ w.r.t. t can be written as follows:

$$q'_t(t, s) = \frac{(\psi_1(t) - \psi_2(t)) h(s) + \psi_2(t) g(s)}{K(t)^2},$$

where

$$\psi_1(t) = D'(t) K(t) - D(t) K'(t)$$

and

$$\psi_2(t) = N'(t) K(t) - N(t) K'(t).$$

Returning to initial notation we obtain the following expression:

$$\begin{aligned} \psi_1(t) = & g + 2I_4g - 3I_2g + I_5hg + I_1h - I_2h + I_1I_4I_5hg + I_4^2g + 3I_2^2g \\ & - I_2I_5h^2 - 2I_2I_5hg - 2I_1I_2h + 2I_2^2h + I_2^2I_3g^2 - I_1I_2I_3h^2 - I_2^3g \\ & + I_2^2I_5hg + I_1I_2^2h - I_2^3h - 2I_2I_4I_5hg + I_1^2I_3h^2 - I_1I_2I_5h^2 - I_3hg \\ & - I_2I_4h + I_3I_4g^2 - I_3I_4hg - I_1I_4g - I_1I_4^2g + I_1I_2I_4g - 2I_2I_3g^2 \\ & + I_1I_2I_4h + I_1I_5h^2 + I_4I_5g^2 + I_4^2g + I_2^2I_3hg - I_1I_3h^2 + I_4^2I_5g^2 \\ & - I_2I_4I_5g^2 + 2I_1I_3hg - 2I_1I_2I_3hg + I_2^2I_4g - I_1^2I_4h + I_2I_3h^2 \\ & - I_2I_3I_4g^2 - 3I_2I_4g + I_2^2I_5h^2 + I_1I_4h + I_1I_3I_4hg + I_4I_5hg + I_3g^2. \end{aligned}$$

The function $\psi_2(t)$ is defined by (24). Hence we obtain

$$\frac{q'_t(t, s)}{q(t, t)} = \frac{(\psi_1(t) - \psi_2(t)) h(s) + \psi_2(t) g(s)}{K(t)^2 + \Phi_1(t)}.$$

Then if we put

$$\varphi_1(t) = \psi_1(t) - \psi_2(t)$$

and

$$\varphi_2(t) = K(t)^2 + \Phi_1(t),$$

then this implies that

$$\frac{q'_t(t, s)}{q(t, t)} = \frac{\varphi_1(t) h(s) + \psi_2(t) g(s)}{\varphi_2(t)},$$

with $\varphi_1(t)$ and $\varphi_2(t)$ were defined by (22)–(23). Finally, we obtain the expression

$$w_t = \int_0^t q(s, s)^{-1} dM_s = U(t) + \int_0^t \int_0^s \frac{\varphi_1(s) h(v) + \psi_2(s) g(v)}{\varphi_2(s)} dU(v) ds.$$

This is the linear transformation $w_t = L[U](t)$ of the process $U(\cdot)$ into the Wiener process w_t and this proves the Theorem 1.

Remark. Let us present a sufficient condition for $\varphi_2(t) > 0$.

\mathcal{R}_0 . Suppose that $h(t)$ and $g(t)$ are continuous strictly positive functions such that $g(t) > h(t)$ and

$$\int_0^t h(s) g(s) ds < 1, \quad \int_0^1 g(s)^2 ds = 1, \quad 0 \leq t < 1,$$

then $\varphi_2(t)$ defined by (23) is strictly positive function on $[0, 1)$.

Now we will verify that if the condition \mathcal{R}_0 is satisfied, then $\varphi_2(t)$ is strictly positive function. Remind that $\varphi_2(t)$ has the following expression:

$$\varphi_2(t) = K(t)^2 + \Phi_1(t), \quad 0 \leq t < 1,$$

where

$$\Phi_1(t) = (D(t) h(t) + N(t) (g(t) - h(t))) K(t),$$

with

$$D(t) = \int_0^t g(s) ds \left(1 + \int_0^t h(s)^2 ds - \int_0^t h(s) g(s) ds \right) + \int_0^t h(s) ds \left(\int_0^t g(s)^2 ds - \int_0^t h(s) g(s) ds \right),$$

$$N(t) = C(t) \int_0^t h(s) ds + \int_0^t h(s)^2 ds \int_0^t g(s) ds$$

and

$$K(t) = C(t)^2 + \int_0^t h(s)^2 ds \int_t^1 g(s)^2 ds.$$

Here

$$C(t) = 1 - \int_0^t h(s) g(s) ds.$$

Note that it is sufficient to check that $\Phi_1(t) > 0$ to obtain $\varphi_2(t) > 0$. Recall that $K(t)$ is strictly positive function. Then due to the following condition

$$\int_0^t h(s) g(s) ds < 1, \quad (40)$$

we have $C(t) > 0$. Consequently, we obtain $N(t) > 0$ by the conditions

$$h(t) > 0 \quad \text{and} \quad g(t) > 0. \quad (41)$$

Finally, we see that $D(t) > 0$ by the conditions $\int_0^1 g(s)^2 ds = 1$, (40) and (41). We conclude that we have $\varphi_2(t) > 0$ if we suppose that $g(t) > h(t)$.

5 Test

Our objective is to test the composite parametric hypothesis \mathcal{H}_0 and to do this we will propose a statistic based on the MDE ϑ_ε^* . Recall that the starting statistic

$$u_\varepsilon(t) = \frac{X_t - x_t(\vartheta_\varepsilon^*)}{\varepsilon S(\vartheta_\varepsilon^*, X_t)} \quad (42)$$

converges to the random function

$$u(t) = \int_0^t \frac{dW_s}{S(\vartheta, x_s)} - \int_0^T \int_v^T \frac{S(\vartheta, x_s) \dot{x}_s(\vartheta)}{J(\vartheta) S(\vartheta, x_v)} ds dW_v \int_0^t \frac{\dot{S}(\vartheta, x_s)}{S(\vartheta, x_s)} ds.$$

Then the linear transformation

$$U\left(\frac{t}{T}\right) = \frac{1}{\sqrt{T}} \int_0^t S(\vartheta, x_s) du(s)$$

has the following representation, by Itô formula,

$$U\left(\frac{t}{T}\right) = \frac{1}{\sqrt{T}} S(\vartheta, x_t) u(t) - \frac{1}{\sqrt{T}} \int_0^t S'(\vartheta, x_s) S(\vartheta, x_s) u(s) ds$$

which leads to the random function

$$U(\nu) = W(\nu) - \int_0^1 g(\vartheta, r) dW(r) \int_0^\nu h(\vartheta, r) dr.$$

The last step is to apply the transformation $L[U](\cdot)$ from Theorem 1 and to obtain the Wiener process

$$L[U](\nu) = w_\nu, \quad 0 \leq \nu \leq 1.$$

Now we have to realize the similar transformations with the “empirical” process $u_\varepsilon(\cdot)$ defined by (42), i.e., we (formally) calculate

$$U_\varepsilon\left(\frac{t}{T}\right) = \frac{1}{\sqrt{T}} \int_0^t S(\vartheta_\varepsilon^*, X_s) du_\varepsilon(s). \quad (43)$$

Then we apply the transformation $L[\cdot]$ to the process $U_\varepsilon(\cdot)$ and we show that this statistic converges in distribution to the Wiener process $w_\nu, 0 \leq \nu \leq 1$. Therefore the test $\psi_\varepsilon = \mathbb{I}_{\{\Delta_\varepsilon > c_\alpha\}}$ with

$$\Delta_\varepsilon = \frac{1}{T} \int_0^T L[U_\varepsilon](t)^2 dt \implies \int_0^1 w_\nu^2 d\nu$$

will be ADF because the limit distribution of Δ_ε does not depend on $S(\cdot, \cdot)$ and ϑ .

Let us realize this program. We have the following representation for the process $U_\varepsilon(\cdot)$

$$\begin{aligned} U_\varepsilon\left(\frac{t}{T}\right) &= \frac{1}{\sqrt{T}} S(\vartheta_\varepsilon^*, X_t) u_\varepsilon(t) \\ &\quad - \frac{1}{\sqrt{T}} \int_0^t S'(\vartheta_\varepsilon^*, X_s) S(\vartheta_\varepsilon^*, X_s) u_\varepsilon(s) ds. \end{aligned} \quad (44)$$

Introduce the functions

$$\hat{h}(\vartheta, v) = \frac{1}{\sqrt{T}} J(\vartheta)^{-1} \dot{S}(\vartheta, x_v), \quad J(\vartheta) = \int_0^T \dot{x}_s(\vartheta)^2 ds \quad (45)$$

and

$$\hat{g}(\vartheta, v) = S(\vartheta, x_v)^{-1} I(\vartheta, v), \quad I(\vartheta, v) = \int_v^T S(\vartheta, x_s) \dot{x}_s(\vartheta) ds \quad (46)$$

and their “empirical versions”, respectively

$$h_\varepsilon(\vartheta_\varepsilon^*, v) = \frac{1}{\sqrt{T}} J_\varepsilon(\vartheta_\varepsilon^*)^{-1} \dot{S}(\vartheta_\varepsilon^*, X_v), \quad g_\varepsilon(\vartheta_\varepsilon^*, v) = S(\vartheta_\varepsilon^*, X_v)^{-1} I_\varepsilon(\vartheta_\varepsilon^*, v).$$

Here

$$J_\varepsilon(\vartheta_\varepsilon^*) = \int_0^T \dot{x}_s(\vartheta_\varepsilon^*)^2 ds, \quad I_\varepsilon(\vartheta_\varepsilon^*, v) = \int_v^T S(\vartheta_\varepsilon^*, X_s) \dot{x}_s(\vartheta_\varepsilon^*) ds.$$

Note that in the functions $\hat{h}(\cdot, \cdot)$ and $\hat{g}(\cdot, \cdot)$ we omit the normalizing constants in the expressions of the functions $h(\cdot, \cdot)$ and $g(\cdot, \cdot)$ defined by (18)–(19), for simplicity of exposition, because the structure of the used statistic is such that we can do this without changing the limit distribution of the statistic.

Then denote the “empirical versions”

$$I_{1,\varepsilon} = C_1(T) \int_0^s g_\varepsilon(\vartheta_\varepsilon^*, v)^2 dv, \quad I_{2,\varepsilon} = C_2(T) \int_0^s h_\varepsilon(\vartheta_\varepsilon^*, v) g_\varepsilon(\vartheta_\varepsilon^*, v) dv,$$

$$I_{3,\varepsilon} = C_3(T) \int_0^s h_\varepsilon(\vartheta_\varepsilon^*, v) \, dv, \quad I_{4,\varepsilon} = C_4(T) \int_0^s h_\varepsilon(\vartheta_\varepsilon^*, v)^2 \, dv$$

and

$$I_{5,\varepsilon} = C_5(T) \int_0^s g_\varepsilon(\vartheta_\varepsilon^*, v) \, dv$$

of the integrals

$$\hat{I}_1 = C_1(T) \int_0^s \hat{g}(\vartheta, v)^2 \, dv, \quad \hat{I}_2 = C_2(T) \int_0^s \hat{h}(\vartheta, v) \hat{g}(\vartheta, v) \, dv,$$

$$\hat{I}_3 = C_3(T) \int_0^s \hat{h}(\vartheta, v) \, dv, \quad \hat{I}_4 = C_4(T) \int_0^s \hat{h}(\vartheta, v)^2 \, dv$$

and

$$\hat{I}_5 = C_5(T) \int_0^s \hat{g}(\vartheta, v) \, dv,$$

where $C_1(T) = \frac{1}{T^3}$, $C_2(T) = \sqrt{T}$, $C_3(T) = T\sqrt{T}$, $C_4(T) = T^4$ and $C_5(T) = \frac{1}{T^2}$. This allows us to introduce the “empirical versions” $\varphi_{1,\varepsilon}(\cdot)$, $\varphi_{2,\varepsilon}(\cdot)$ and $\psi_{2,\varepsilon}(\cdot)$ of $\hat{\varphi}_1(\cdot)$, $\hat{\varphi}_2(\cdot)$ and $\hat{\psi}_2(\cdot)$ defined respectively by (22)–(24), where we replace the functions g by $\hat{g} = \frac{1}{T} \hat{g}(\vartheta, s)$ and h by $\hat{h} = T^2 \sqrt{T} \hat{h}(\vartheta, s)$,

$$\begin{aligned} \varphi_{1,\varepsilon}(s) &= g_\varepsilon - h_\varepsilon + I_{4,\varepsilon} g_\varepsilon - 3I_{2,\varepsilon} g_\varepsilon + I_{5,\varepsilon} h_\varepsilon g_\varepsilon + I_{1,\varepsilon} h_\varepsilon + 2I_{2,\varepsilon} h_\varepsilon + I_{2,\varepsilon}^2 I_{3,\varepsilon} g_\varepsilon^2 \\ &\quad - 2I_{2,\varepsilon} I_{3,\varepsilon} g_\varepsilon^2 - I_{5,\varepsilon} h_\varepsilon^2 + I_{3,\varepsilon} g_\varepsilon^2 - I_{2,\varepsilon} I_{4,\varepsilon} g_\varepsilon + 3I_{2,\varepsilon}^2 g_\varepsilon - 2I_{3,\varepsilon} h_\varepsilon g_\varepsilon \\ &\quad - 2I_{2,\varepsilon} I_{5,\varepsilon} h_\varepsilon g_\varepsilon - I_{2,\varepsilon}^3 g_\varepsilon + I_{2,\varepsilon}^2 I_{5,\varepsilon} h_\varepsilon g_\varepsilon + I_{1,\varepsilon} I_{2,\varepsilon}^2 h_\varepsilon - I_{2,\varepsilon}^2 h_\varepsilon \\ &\quad + 2I_{1,\varepsilon} I_{4,\varepsilon} h_\varepsilon - I_{1,\varepsilon} I_{4,\varepsilon} g_\varepsilon + I_{1,\varepsilon}^2 I_{3,\varepsilon} h_\varepsilon^2 + I_{1,\varepsilon} I_{4,\varepsilon} I_{5,\varepsilon} h_\varepsilon g_\varepsilon + I_{2,\varepsilon} I_{5,\varepsilon} h_\varepsilon^2 \\ &\quad + I_{1,\varepsilon} I_{5,\varepsilon} h_\varepsilon^2 + I_{4,\varepsilon} I_{5,\varepsilon} g_\varepsilon^2 - I_{1,\varepsilon} I_{2,\varepsilon} I_{5,\varepsilon} h_\varepsilon^2 - I_{4,\varepsilon} h_\varepsilon - I_{2,\varepsilon} I_{4,\varepsilon} I_{5,\varepsilon} g_\varepsilon^2 \\ &\quad + 2I_{1,\varepsilon} I_{3,\varepsilon} h_\varepsilon g_\varepsilon - 2I_{1,\varepsilon} I_{2,\varepsilon} I_{3,\varepsilon} h_\varepsilon g_\varepsilon - 2I_{1,\varepsilon} I_{3,\varepsilon} h_\varepsilon^2 - 2I_{1,\varepsilon} I_{2,\varepsilon} h_\varepsilon \\ &\quad + I_{1,\varepsilon} I_{2,\varepsilon} I_{4,\varepsilon} g_\varepsilon + 2I_{2,\varepsilon} I_{3,\varepsilon} h_\varepsilon g_\varepsilon - I_{4,\varepsilon} I_{5,\varepsilon} h_\varepsilon g_\varepsilon + I_{3,\varepsilon} h_\varepsilon^2 - I_{1,\varepsilon}^2 I_{4,\varepsilon} h_\varepsilon, \\ \varphi_{2,\varepsilon}(s) &= 1 - 2I_{1,\varepsilon} I_{4,\varepsilon}^2 - 3I_{2,\varepsilon} I_{5,\varepsilon} h_\varepsilon + I_{1,\varepsilon} I_{3,\varepsilon} h_\varepsilon + I_{3,\varepsilon} g_\varepsilon - 3I_{2,\varepsilon} I_{3,\varepsilon} g_\varepsilon \\ &\quad + I_{4,\varepsilon} I_{5,\varepsilon} g_\varepsilon - I_{3,\varepsilon} h_\varepsilon + 3I_{2,\varepsilon}^2 I_{5,\varepsilon} h_\varepsilon - 2I_{1,\varepsilon} I_{2,\varepsilon} I_{3,\varepsilon} h_\varepsilon - 2I_{1,\varepsilon} I_{2,\varepsilon}^2 I_{4,\varepsilon} \\ &\quad - 2I_{2,\varepsilon} I_{4,\varepsilon} I_{5,\varepsilon} g_\varepsilon + 2I_{2,\varepsilon} I_{3,\varepsilon} h_\varepsilon - I_{2,\varepsilon}^3 I_{5,\varepsilon} h_\varepsilon + I_{1,\varepsilon} I_{2,\varepsilon}^2 I_{3,\varepsilon} h_\varepsilon - 2I_{1,\varepsilon} I_{4,\varepsilon} \\ &\quad - I_{2,\varepsilon}^3 I_{3,\varepsilon} g_\varepsilon + I_{4,\varepsilon}^2 + I_{2,\varepsilon}^2 I_{4,\varepsilon} I_{5,\varepsilon} g_\varepsilon + I_{4,\varepsilon} I_{5,\varepsilon} h_\varepsilon + I_{4,\varepsilon}^2 I_{5,\varepsilon} g_\varepsilon \\ &\quad - I_{2,\varepsilon} I_{4,\varepsilon} I_{5,\varepsilon} h_\varepsilon + 2I_{1,\varepsilon} I_{3,\varepsilon} I_{4,\varepsilon} h_\varepsilon + I_{3,\varepsilon} I_{4,\varepsilon} g_\varepsilon + 3I_{2,\varepsilon}^2 I_{3,\varepsilon} g_\varepsilon \\ &\quad + 2I_{2,\varepsilon}^2 I_{4,\varepsilon} + 2I_{4,\varepsilon} - I_{3,\varepsilon} I_{4,\varepsilon} h_\varepsilon + 6I_{2,\varepsilon}^2 + I_{1,\varepsilon} I_{2,\varepsilon} I_{4,\varepsilon} I_{5,\varepsilon} h_\varepsilon \\ &\quad - I_{1,\varepsilon}^2 I_{3,\varepsilon} I_{4,\varepsilon} h_\varepsilon - I_{1,\varepsilon} I_{3,\varepsilon} I_{4,\varepsilon} g_\varepsilon + I_{5,\varepsilon} h_\varepsilon + I_{1,\varepsilon} I_{2,\varepsilon} I_{3,\varepsilon} I_{4,\varepsilon} g_\varepsilon \\ &\quad - I_{1,\varepsilon} I_{4,\varepsilon}^2 I_{5,\varepsilon} g_\varepsilon - I_{1,\varepsilon} I_{4,\varepsilon} I_{5,\varepsilon} h_\varepsilon + 4I_{1,\varepsilon} I_{2,\varepsilon} I_{4,\varepsilon} - I_{2,\varepsilon} I_{3,\varepsilon} I_{4,\varepsilon} g_\varepsilon \end{aligned}$$

$$- I_{2,\varepsilon}^2 I_{3,\varepsilon} h_\varepsilon + I_{2,\varepsilon}^4 - 4 I_{2,\varepsilon}^3 - 4 I_{2,\varepsilon} - 4 I_{2,\varepsilon} I_{4,\varepsilon} + I_{1,\varepsilon}^2 I_{4,\varepsilon}^2 \quad (47)$$

and

$$\begin{aligned} \psi_{2,\varepsilon}(s) = & h_\varepsilon + I_{3,\varepsilon} h_\varepsilon g_\varepsilon + I_{4,\varepsilon} g_\varepsilon + I_{5,\varepsilon} h_\varepsilon^2 - 3 I_{2,\varepsilon} h_\varepsilon - I_{1,\varepsilon} I_{2,\varepsilon} I_{3,\varepsilon} h_\varepsilon^2 \\ & - 2 I_{2,\varepsilon} I_{4,\varepsilon} g_\varepsilon + I_{4,\varepsilon} h_\varepsilon - 2 I_{2,\varepsilon} I_{5,\varepsilon} h_\varepsilon^2 + 3 I_{2,\varepsilon}^2 h_\varepsilon + I_{2,\varepsilon}^2 I_{4,\varepsilon} g_\varepsilon \\ & + I_{2,\varepsilon}^2 I_{5,\varepsilon} h_\varepsilon^2 - I_{2,\varepsilon}^3 h_\varepsilon - I_{3,\varepsilon} I_{4,\varepsilon} h_\varepsilon g_\varepsilon + I_{4,\varepsilon}^2 g_\varepsilon - I_{2,\varepsilon} I_{4,\varepsilon} h_\varepsilon \\ & + I_{1,\varepsilon} I_{3,\varepsilon} I_{4,\varepsilon} h_\varepsilon g_\varepsilon - I_{1,\varepsilon} I_{4,\varepsilon}^2 g_\varepsilon - I_{1,\varepsilon} I_{4,\varepsilon} h_\varepsilon + I_{1,\varepsilon} I_{2,\varepsilon} I_{4,\varepsilon} h_\varepsilon \\ & - I_{3,\varepsilon} h_\varepsilon^2 + I_{1,\varepsilon} I_{3,\varepsilon} h_\varepsilon^2 + I_{3,\varepsilon} I_{4,\varepsilon} g_\varepsilon^2 + I_{2,\varepsilon} I_{3,\varepsilon} h_\varepsilon^2 - 2 I_{2,\varepsilon} I_{3,\varepsilon} h_\varepsilon g_\varepsilon \\ & - 2 I_{2,\varepsilon} I_{4,\varepsilon} I_{5,\varepsilon} h_\varepsilon g_\varepsilon + 2 I_{4,\varepsilon} I_{5,\varepsilon} h_\varepsilon g_\varepsilon - I_{2,\varepsilon} I_{3,\varepsilon} I_{4,\varepsilon} g_\varepsilon^2 + I_{2,\varepsilon}^2 I_{3,\varepsilon} h_\varepsilon g_\varepsilon \\ & + I_{4,\varepsilon}^2 I_{5,\varepsilon} g_\varepsilon^2. \end{aligned}$$

Here $g_\varepsilon = g_\varepsilon(\vartheta_\varepsilon^*, s)$, $h_\varepsilon = h_\varepsilon(\vartheta_\varepsilon^*, s)$ and $I_{1,\varepsilon}$, $I_{2,\varepsilon}$, $I_{3,\varepsilon}$, $I_{4,\varepsilon}$, $I_{5,\varepsilon}$ are the “empirical versions” of \hat{I}_1 , \hat{I}_2 , \hat{I}_3 , \hat{I}_4 , \hat{I}_5 , respectively.

In the construction of the test we introduce one condition else.

\mathcal{R}_1 . We suppose that $\varphi_2(r)$, $r \in [0, 1)$ defined by (23) is strictly positive function.

We have the uniform convergence in probability w.r.t. $s \in [0, T]$ ($\varepsilon \rightarrow 0$)

$$\sup_{s \in [0, T]} |\varphi_{2,\varepsilon}(s) - \hat{\varphi}_2(s)| \longrightarrow 0.$$

This convergence we obtain due to the consistency of the estimator and the smoothness of the functions $g_\varepsilon(\cdot, \cdot)$ and $h_\varepsilon(\cdot, \cdot)$.

Therefore we can introduce the function

$$\varphi_{2,\varepsilon}^+(s) = \begin{cases} \varphi_{2,\varepsilon}(s)^{-1}, & \text{if } \varphi_{2,\varepsilon}(s) > 0, \\ 0, & \text{else,} \end{cases}$$

which asymptotically coincides with $\hat{\varphi}_2(s)^{-1}$ and therefore the limit distribution does not change.

Hence we consider (formally) the statistic

$$\begin{aligned} W_\varepsilon(t) = & U_\varepsilon\left(\frac{t}{T}\right) \\ & + \frac{1}{T} \int_0^t \int_0^s \varphi_{2,\varepsilon}^+(s) [\lambda_1(T) \varphi_{1,\varepsilon}(s) h_\varepsilon(\vartheta_\varepsilon^*, q) + \lambda_2(T) \psi_{2,\varepsilon}(s) g_\varepsilon(\vartheta_\varepsilon^*, q)] dU_\varepsilon\left(\frac{q}{T}\right) ds \end{aligned} \quad (48)$$

where $\lambda_1(T) = T^2 \sqrt{T}$, $\lambda_2(T) = \frac{1}{T}$ and $U_\varepsilon(\cdot)$ was defined by (44). If we prove that

$$W_\varepsilon(t) \longrightarrow L[U](\nu) = w_\nu,$$

then the test based on this statistic will be ADF. The main technical problem in carrying out this program is to define the stochastic integrals

$$K_\varepsilon(\vartheta_\varepsilon^*, s) = \int_0^s h_\varepsilon(\vartheta_\varepsilon^*, q) \, dU_\varepsilon\left(\frac{q}{T}\right) \quad (49)$$

and

$$L_\varepsilon(\vartheta_\varepsilon^*, s) = \int_0^s g_\varepsilon(\vartheta_\varepsilon^*, q) \, dU_\varepsilon\left(\frac{q}{T}\right). \quad (50)$$

Unfortunately we can not calculate them as they are written now, because the integrand contains the MDE ϑ_ε^* and this estimator depends on the whole trajectory $X^\varepsilon = (X_t, 0 \leq t \leq T)$. Therefore the corresponding stochastic integrals $K_\varepsilon(\vartheta_\varepsilon^*, s)$ and $L_\varepsilon(\vartheta_\varepsilon^*, s)$ are not well defined.

To avoid this problem we use an approach which is based on the application of the Itô formula, i.e., we replace the corresponding stochastic integrals by the ordinary ones. Note that this approach was applied in the similar problem in [12].

Introduce the statistic

$$K(\vartheta, s) = \int_0^s \hat{h}(\vartheta, q) \, dU\left(\frac{q}{T}\right) = \frac{1}{\sqrt{T}} \int_0^s \hat{h}(\vartheta, q) S(\vartheta, x_q) \, du(q), \quad (51)$$

where the process $U(\cdot)$ and the function $\hat{h}(\cdot, \cdot)$ were defined by (15) and (45), respectively. Indeed the Itô formula gives us the following representation

$$\begin{aligned} d(\hat{h}(\vartheta, q) S(\vartheta, x_q) u(q)) &= \left(\hat{h}'(\vartheta, q) S(\vartheta, x_q) + \hat{h}(\vartheta, q) S'(\vartheta, x_q) S(\vartheta, x_q) \right) u(q) dq \\ &\quad + \hat{h}(\vartheta, q) S(\vartheta, x_q) \, du(q). \end{aligned}$$

Here $\hat{h}'(\vartheta, q)$ is the derivative of $\hat{h}(\vartheta, q)$ w.r.t. q , given by the following expression

$$\hat{h}'(\vartheta, q) = \frac{1}{\sqrt{T}} J(\vartheta)^{-1} \dot{S}'(\vartheta, x_q(\vartheta)) S(\vartheta, x_q(\vartheta)). \quad (52)$$

Therefore the statistic $K(\vartheta, \cdot)$ defined by (51) can be written as follows:

$$\begin{aligned} K(\vartheta, s) &= \frac{1}{\sqrt{T}} \hat{h}(\vartheta, s) S(\vartheta, x_s) u(s) \\ &\quad - \frac{1}{\sqrt{T}} \int_0^s \left(\hat{h}'(\vartheta, q) S(\vartheta, x_q) + \hat{h}(\vartheta, q) S'(\vartheta, x_q) S(\vartheta, x_q) \right) u(q) \, dq. \end{aligned}$$

Hence we obtain for the process (49) the following representation

$$\begin{aligned} K_\varepsilon(\vartheta_\varepsilon^*, s) &= \frac{1}{\sqrt{T}} h_\varepsilon(\vartheta_\varepsilon^*, s) S(\vartheta_\varepsilon^*, X_s) u_\varepsilon(s) \\ &- \frac{1}{\sqrt{T}} \int_0^s (h'_\varepsilon(\vartheta_\varepsilon^*, q) S(\vartheta_\varepsilon^*, X_q) + h_\varepsilon(\vartheta_\varepsilon^*, q) S'(\vartheta_\varepsilon^*, X_q) S(\vartheta_\varepsilon^*, X_q)) u_\varepsilon(q) dq \end{aligned} \quad (53)$$

and the integral is now well defined. Here

$$h'_\varepsilon(\vartheta_\varepsilon^*, q) = \frac{1}{\sqrt{T}} J_\varepsilon(\vartheta_\varepsilon^*)^{-1} \dot{S}'(\vartheta_\varepsilon^*, X_q) S(\vartheta_\varepsilon^*, X_q).$$

Similarly, the process (50) can be written as follows:

$$\begin{aligned} L_\varepsilon(\vartheta_\varepsilon^*, s) &= \frac{1}{\sqrt{T}} g_\varepsilon(\vartheta_\varepsilon^*, s) S(\vartheta_\varepsilon^*, X_s) u_\varepsilon(s) \\ &- \frac{1}{\sqrt{T}} \int_0^s (g'_\varepsilon(\vartheta_\varepsilon^*, q) S(\vartheta_\varepsilon^*, X_q) + g_\varepsilon(\vartheta_\varepsilon^*, q) S'(\vartheta_\varepsilon^*, X_q) S(\vartheta_\varepsilon^*, X_q)) u_\varepsilon(q) dq, \end{aligned} \quad (54)$$

where

$$g'_\varepsilon(\vartheta_\varepsilon^*, q) = -S(\vartheta_\varepsilon^*, X_q)^{-1} S'(\vartheta_\varepsilon^*, X_q) \int_q^T S(\vartheta_\varepsilon^*, X_s) \dot{x}_s(\vartheta_\varepsilon^*) ds - \dot{x}_q(\vartheta_\varepsilon^*)$$

is the “empirical version” of

$$\hat{g}'(\vartheta, q) = -S(\vartheta, x_q(\vartheta))^{-1} S'(\vartheta, x_q(\vartheta)) \int_q^T S(\vartheta, x_s(\vartheta)) \dot{x}_s(\vartheta) ds - \dot{x}_q(\vartheta). \quad (55)$$

Then the formal expression (48) for $W_\varepsilon(t)$ can be replaced by

$$\begin{aligned} \tilde{W}_\varepsilon(t) &= U_\varepsilon \left(\frac{t}{T} \right) \\ &+ \frac{1}{T} \int_0^t \varphi_{2,\varepsilon}^+(s) [\lambda_1(T) \varphi_{1,\varepsilon}(s) K_\varepsilon(\vartheta_\varepsilon^*, s) + \lambda_2(T) \psi_{2,\varepsilon}(s) L_\varepsilon(\vartheta_\varepsilon^*, s)] ds, \end{aligned} \quad (56)$$

where $\lambda_1(T) = T^2 \sqrt{T}$, $\lambda_2(T) = \frac{1}{T}$ and the processes $U_\varepsilon(\cdot)$, $K_\varepsilon(\vartheta_\varepsilon^*, \cdot)$ and $L_\varepsilon(\vartheta_\varepsilon^*, \cdot)$ admit the representations (44), (53) and (54), respectively.

The test is given in the following Theorem.

Theorem 2 *Suppose that the conditions of regularity \mathcal{R} and \mathcal{R}_1 are fulfilled, then the test*

$$\psi_\varepsilon = \mathbb{I}_{\{\Delta_\varepsilon > c_\alpha\}}, \quad \mathbf{P}(\tilde{\Delta} > c_\alpha) = \alpha,$$

with

$$\Delta_\varepsilon = \frac{1}{T} \int_0^T \tilde{W}_\varepsilon(t)^2 dt, \quad \tilde{\Delta} \equiv \int_0^1 w_\nu^2 d\nu$$

is ADF and of asymptotic size $\alpha \in (0, 1)$.

Proof. We have to show that, under hypothesis \mathcal{H}_0 , the convergence

$$\Delta_\varepsilon \implies \tilde{\Delta} \quad (57)$$

holds.

Recall that the process $U_\varepsilon(\cdot)$ has the following representation

$$\begin{aligned} U_\varepsilon\left(\frac{t}{T}\right) &= \frac{1}{\sqrt{T}} S(\vartheta_\varepsilon^*, X_t) u_\varepsilon(t) \\ &\quad - \frac{1}{\sqrt{T}} \int_0^t S'(\vartheta_\varepsilon^*, X_s) S(\vartheta_\varepsilon^*, X_s) u_\varepsilon(s) ds. \end{aligned}$$

Note that we have already the convergence in probability uniformly w.r.t. $s \in [0, T]$ (as $\varepsilon \rightarrow 0$)

$$\sup_{s \in [0, T]} |X_s - x_s(\vartheta)| \longrightarrow 0, \quad \sup_{s \in [0, T]} |u_\varepsilon(s) - u(s)| \longrightarrow 0.$$

Further, we can write

$$\begin{aligned} |S(\vartheta_\varepsilon^*, X_s) - S(\vartheta, x_s)| &\leq |S(\vartheta_\varepsilon^*, X_s) - S(\vartheta, X_s)| + |S(\vartheta, X_s) - S(\vartheta, x_s)| \\ &\leq |\vartheta_\varepsilon^* - \vartheta| \left| \dot{S}(\tilde{\vartheta}, X_s) \right| + |X_s - x_s| \left| S'(\vartheta, \tilde{X}_s) \right|. \end{aligned}$$

Here $|\tilde{\vartheta} - \vartheta| \leq |\vartheta_\varepsilon^* - \vartheta|$ and

$$\left| \tilde{X}_s - X_s \right| \leq |x_s(\vartheta_\varepsilon^*) - X_s| \leq |x_s(\vartheta_\varepsilon^*) - x_s(\vartheta)| + |x_s(\vartheta) - X_s| \rightarrow 0.$$

This convergence is uniform w.r.t. $s \in [0, T]$.

Then we know that the functions $\dot{S}(\vartheta, x)$ and $S'(\vartheta, x)$ are bounded by regularity conditions \mathcal{R} , the process X_s converges uniformly w.r.t. $s \in [0, T]$ to $x_s(\vartheta)$ and due to the consistency of the estimator ϑ_ε^* , we obtain the uniform convergence w.r.t. $s \in [0, T]$ (in probability)

$$\sup_{s \in [0, T]} |S(\vartheta_\varepsilon^*, X_s) - S(\vartheta, x_s)| \longrightarrow 0.$$

Further, similar arguments give the uniform convergence w.r.t. $s \in [0, T]$ (in probability)

$$\sup_{s \in [0, T]} |S'(\vartheta_\varepsilon^*, X_s) - S'(\vartheta, x_s)| \longrightarrow 0.$$

Therefore, we obtain the uniform convergence w.r.t. $t \in [0, T]$ (in probability)

$$\begin{aligned} U_\varepsilon\left(\frac{t}{T}\right) &\longrightarrow \frac{1}{\sqrt{T}} S(\vartheta, x_t) u(t) - \frac{1}{\sqrt{T}} \int_0^t S'(\vartheta, x_s) S(\vartheta, x_s) u(s) \, ds \\ &= \frac{1}{\sqrt{T}} \int_0^t S(\vartheta, x_s) \, du(s) = U\left(\frac{t}{T}\right). \end{aligned}$$

Now we have to show that $K_\varepsilon(\vartheta_\varepsilon^*, s) \longrightarrow K(\vartheta, s)$, where

$$\begin{aligned} K_\varepsilon(\vartheta_\varepsilon^*, s) &= \frac{1}{\sqrt{T}} h_\varepsilon(\vartheta_\varepsilon^*, s) S(\vartheta_\varepsilon^*, X_s) u_\varepsilon(s) \\ &\quad - \frac{1}{\sqrt{T}} \int_0^s (h'_\varepsilon(\vartheta_\varepsilon^*, q) S(\vartheta_\varepsilon^*, X_q) + h_\varepsilon(\vartheta_\varepsilon^*, q) S'(\vartheta_\varepsilon^*, X_q) S(\vartheta_\varepsilon^*, X_q)) u_\varepsilon(q) \, dq \end{aligned}$$

and

$$\begin{aligned} K(\vartheta, s) &= \frac{1}{\sqrt{T}} \hat{h}(\vartheta, s) S(\vartheta, x_s) u(s) \\ &\quad - \frac{1}{\sqrt{T}} \int_0^s \left(\hat{h}'(\vartheta, q) S(\vartheta, x_q) + \hat{h}(\vartheta, q) S'(\vartheta, x_q) S(\vartheta, x_q) \right) u(q) \, dq, \end{aligned}$$

where $\hat{h}(\cdot, \cdot)$ and $\hat{h}'(\cdot, \cdot)$ were defined by (45)–(52).

We have

$$\begin{aligned} \left| \dot{S}(\vartheta_\varepsilon^*, X_s) - \dot{S}(\vartheta, x_s) \right| &\leq \left| \dot{S}(\vartheta_\varepsilon^*, X_s) - \dot{S}(\vartheta, X_s) \right| + \left| \dot{S}(\vartheta, X_s) - \dot{S}(\vartheta, x_s) \right| \\ &\leq |\vartheta_\varepsilon^* - \vartheta| \left| \ddot{S}(\vartheta, X_s) \right| + |X_s - x_s| \left| \dot{S}'(\vartheta, \tilde{X}_s) \right| \\ &\leq \tilde{C}_1 |\vartheta_\varepsilon^* - \vartheta| + \tilde{C}_2 |X_s - x_s|. \end{aligned}$$

Here \tilde{C}_1 and \tilde{C}_2 are some constants and $\ddot{S}(\vartheta, x)$, which means the second derivative w.r.t. ϑ , and $\dot{S}'(\vartheta, x)$ are bounded functions due to the regularity conditions \mathcal{R} . Therefore we have the convergence in probability

$$\sup_{s \in [0, T]} \left| \dot{S}(\vartheta_\varepsilon^*, X_s) - \dot{S}(\vartheta, x_s) \right| \longrightarrow 0.$$

Further, we have

$$J_\varepsilon(\vartheta_\varepsilon^*) - J(\vartheta) = \int_0^T \dot{x}_s(\vartheta_\varepsilon^*)^2 \, ds - \int_0^T \dot{x}_s(\vartheta)^2 \, ds \longrightarrow 0.$$

Hence we obtain the convergence in probability

$$\sup_{s \in [0, T]} \left| h_\varepsilon(\vartheta_\varepsilon^*, s) - \hat{h}(\vartheta, s) \right| \longrightarrow 0.$$

Similarly, it is shown that we have the convergence in probability

$$\sup_{s \in [0, T]} \left| h'_\varepsilon(\vartheta_\varepsilon^*, s) - \hat{h}'(\vartheta, s) \right| \longrightarrow 0.$$

Therefore we obtain the convergence in probability, by the uniform convergence of $u_\varepsilon(s)$ to $u(s)$ w.r.t. $s \in [0, T]$,

$$K_\varepsilon(\vartheta_\varepsilon^*, s) \longrightarrow K(\vartheta, s).$$

Now we have to show that $L_\varepsilon(\vartheta_\varepsilon^*, s) \longrightarrow L(\vartheta, s)$, where

$$\begin{aligned} L_\varepsilon(\vartheta_\varepsilon^*, s) &= \frac{1}{\sqrt{T}} g_\varepsilon(\vartheta_\varepsilon^*, s) S(\vartheta_\varepsilon^*, X_s) u_\varepsilon(s) \\ &\quad - \frac{1}{\sqrt{T}} \int_0^s (g'_\varepsilon(\vartheta_\varepsilon^*, q) S(\vartheta_\varepsilon^*, X_q) + g_\varepsilon(\vartheta_\varepsilon^*, q) S'(\vartheta_\varepsilon^*, X_q) S(\vartheta_\varepsilon^*, X_q)) u_\varepsilon(q) \, dq \end{aligned}$$

and

$$\begin{aligned} L(\vartheta, s) &= \frac{1}{\sqrt{T}} \hat{g}(\vartheta, s) S(\vartheta, x_s) u(s) \\ &\quad - \frac{1}{\sqrt{T}} \int_0^s (\hat{g}'(\vartheta, q) S(\vartheta, x_q) + \hat{g}(\vartheta, q) S'(\vartheta, x_q) S(\vartheta, x_q)) u(q) \, dq, \end{aligned}$$

where $\hat{g}(\cdot, \cdot)$ and $\hat{g}'(\cdot, \cdot)$ were defined by (46) and (55), respectively.

Observe that for $s \in [0, T]$, we have

$$\begin{aligned} |I_\varepsilon(\vartheta_\varepsilon^*, s) - I(\vartheta, s)| &\leq \int_s^T |S(\vartheta_\varepsilon^*, X_v) (\dot{x}_v(\vartheta_\varepsilon^*) - \dot{x}_v(\vartheta))| \, dv \\ &\quad + \int_s^T |\dot{x}_v(\vartheta) (S(\vartheta_\varepsilon^*, X_v) - S(\vartheta, x_v(\vartheta)))| \, dv \\ &\leq \int_s^T |S(\vartheta_\varepsilon^*, X_v)| |\dot{x}_v(\vartheta_\varepsilon^*) - \dot{x}_v(\vartheta)| \, dv \\ &\quad + |\vartheta_\varepsilon^* - \vartheta| \int_s^T |\dot{x}_v(\vartheta)| \left| \dot{S}(\tilde{\vartheta}, X_v) \right| \, dv \\ &\quad + \int_s^T |\dot{x}_v(\vartheta)| |X_v - x_v(\vartheta)| \left| S'(\vartheta, \tilde{X}_v) \right| \, dv \end{aligned}$$

$$\leq \left(\tilde{C}_3 |\vartheta_\varepsilon^* - \vartheta| + \tilde{C}_4 |\vartheta_\varepsilon^* - \vartheta| + \tilde{C}_5 \sup_{s \in [0, T]} |X_s - x_s(\vartheta)| \right) (T - s),$$

where \tilde{C}_3 , \tilde{C}_4 and \tilde{C}_5 are constants. Therefore we obtain the convergence in probability uniformly w.r.t. $s \in [0, T]$

$$\sup_{s \in [0, T]} |I_\varepsilon(\vartheta_\varepsilon^*, s) - I(\vartheta, s)| \longrightarrow 0$$

because the estimator ϑ_ε^* is consistent, the process X_s converges uniformly w.r.t. $s \in [0, T]$ to $x_s(\vartheta)$ and the derivatives are bounded due to the conditions of regularity \mathcal{R} . Further, we proved already the uniform convergence w.r.t. $s \in [0, T]$ of $u_\varepsilon(s)$ to $u(s)$ and $S(\vartheta_\varepsilon^*, X_s)$ to $S(\vartheta, x_v)$. Hence, we have the convergence in probability

$$\sup_{s \in [0, T]} |g_\varepsilon(\vartheta_\varepsilon^*, s) - \hat{g}(\vartheta, s)| \longrightarrow 0.$$

Similarly, by the regularity conditions \mathcal{R} , it is shown that (in probability)

$$\sup_{s \in [0, T]} |g'_\varepsilon(\vartheta_\varepsilon^*, s) - \hat{g}'(\vartheta, s)| \longrightarrow 0.$$

Therefore we obtain the convergence in probability

$$L_\varepsilon(\vartheta_\varepsilon^*, s) \longrightarrow L(\vartheta, s).$$

Further, a similar arguments give the convergence in probability uniformly w.r.t. $s \in [0, T]$ due to the regularity conditions \mathcal{R} and the consistency of the estimator ϑ_ε^*

$$\sup_{s \in [0, T]} \left| \psi_{2, \varepsilon}(s) - \hat{\psi}_2(s) \right| \longrightarrow 0$$

and

$$\sup_{s \in [0, T]} |\varphi_{1, \varepsilon}(s) - \hat{\varphi}_1(s)| \longrightarrow 0, \quad \sup_{s \in [0, T]} |\varphi_{2, \varepsilon}(s) - \hat{\varphi}_2(s)| \longrightarrow 0.$$

Finally, the convergence mentioned in (57) is proved and using this result the test ψ_ε is ADF and of asymptotic size $\alpha \in (0, 1)$.

6 The case of MLE

This case was studied in [6]–[7]. They proposed a linear transformation, which yields the convergence of the test statistic to the integral of Wiener process. Therefore they showed that the test based on this statistic is ADF.

To obtain the linear transformation mentioned in [7], we put $h(\vartheta, r) = g(\vartheta, r)$ in (21) and we obtain

$$I_1 = I_2 = I_4 = \int_0^r h(\vartheta, q)^2 dq, \quad I_3 = I_5 = \int_0^r h(\vartheta, q) dq.$$

Therefore we can write

$$\begin{aligned} \varphi_1(r) &= h - h + I_1 h - 3I_1 h + I_3 h^2 + I_3 h^2 + 2I_1 h - I_1^2 h - 2I_1 I_3 h^2 - I_1^2 I_3 h^2 \\ &\quad - I_1 h + I_1^2 I_3 h^2 - I_1^2 h + 3I_1^2 h + I_1 I_3 h^2 - 2I_1 I_3 h^2 - 2I_1^2 h + 2I_1 I_3 h^2 \\ &\quad + I_1^2 I_3 h^2 + I_1^3 h + I_3 h^2 + I_1 h + 2I_1^2 h - I_1^2 h + I_1^3 h + I_1^2 I_3 h^2 - 2I_1^2 I_3 h^2 \\ &\quad - I_1^3 h + I_1 I_3 h^2 + I_1 I_3 h^2 + 2I_1 I_3 h^2 - I_1^3 h - I_1^2 I_3 h^2 - I_3 h^2 - 2I_1 I_3 h^2 \\ &\quad + I_1^2 I_3 h^2 - I_1 I_3 h^2 - 2I_3 h^2 \\ &= 0, \\ \varphi_2(r) &= 1 + I_3 h - 3I_1 I_3 h + I_1 I_3 h + I_3 h - 3I_1 I_3 h + I_1 I_3 h - I_3 h + 6I_1^2 - 2I_1^4 \\ &\quad + 3I_1^2 I_3 h - 2I_1^2 I_3 h + 2I_1 I_3 h - I_1^3 I_3 h + I_1^3 I_3 h - I_1^3 I_3 h + I_1^2 - 4I_1 \\ &\quad + I_1 I_3 h - I_1^2 I_3 h + 2I_1^2 I_3 h + I_1 I_3 h - I_1^2 I_3 h + 2I_1^3 + I_1^2 I_3 h - 4I_1^3 \\ &\quad + I_1^3 I_3 h - I_1^3 I_3 h - I_1^2 I_3 h + I_1^3 I_3 h - I_1^2 I_3 h + 4I_1^3 + 3I_1^2 I_3 h - I_1^3 I_3 h \\ &\quad + I_1^4 - 2I_1^3 + I_1^4 + 2I_1 - 2I_1^2 + I_1^3 I_3 h - 2I_1^2 I_3 h - I_1 I_3 h - I_1^2 I_3 h - 4I_1^2 \\ &= 1 + I_3 h - I_1 I_3 h + I_1^2 - 2I_1 \\ &= (1 - I_1)(1 + I_3 h - I_1) \end{aligned}$$

and

$$\begin{aligned} \psi_2(r) &= h + I_3 h^2 + I_1 h + I_3 h^2 - 3I_1 h - 2I_1 I_3 h^2 - 2I_1^2 h - 2I_1 I_3 h^2 - I_1^3 h \\ &\quad + I_1^3 h + I_1^2 I_3 h^2 - I_1^3 h - I_1 I_3 h^2 + I_1^2 h + I_1 h - I_1^2 h - I_3 h^2 + 3I_1^2 h \\ &\quad - I_1^2 h + I_1^3 h + I_1 I_3 h^2 + I_1 I_3 h^2 + I_1 I_3 h^2 - I_1^2 I_3 h^2 + 2I_1^2 I_3 h^2 \\ &\quad - I_1^2 I_3 h^2 - 2I_1^2 I_3 h^2 + 2I_1 I_3 h^2 + I_1^2 I_3 h^2 \\ &= h (1 + I_3 h - I_1), \end{aligned}$$

where $h = h(\vartheta, r)$. Hence the linear transformation (25) will have the following expression:

$$L[U](\nu) = U(\nu) + \int_0^\nu \int_0^r h(r) \mathbb{N}(r)^{-1} h(q) dU(q) dr = w_\nu, \quad (58)$$

where $\mathbb{N}(r) = \int_r^1 h(q)^2 dq$ and

$$U(\nu) = W(\nu) - \int_0^1 h(r) dW(r) \int_0^\nu h(r) dr,$$

with $W(\nu)$ and $w_\nu, 0 \leq \nu \leq 1$ are some standard Wiener processes.

The transformation $L[U](\cdot)$ of the limit process $U(\cdot)$ given by (58) coincides with one by Khmaladze [6].

Acknowledgements. The author is grateful to Yu. A. Kutoyants for the statement of the problem and many helpful suggestions during the preparation of this paper. The author would like to thank the two Referees for the comments which allowed to improve essentially the exposition.

References

- [1] Dachian, S. and Kutoyants, Yu. A. (2008) On the goodness-of-fit tests for some continuous time processes, in *Statistical Models and Methods for Biomedical and Technical Systems*, (F. Vonta, M. Nikulin, N. Limnios and C. Huber-Carol, eds). Boston, 385-403.
- [2] Darling, D. A. (1954) The Kolmogorov-Smirnov, Cramér-von Mises tests. *Ann. Math. Stat.*, 28(4), 823-838.
- [3] Durbin, J. (1973) *Distribution Theory for Tests Based on the Sample Distribution Function*. SIAM, Philadelphia.
- [4] Freidlin M. I. and Wentzell, A. D. (1984) *Random Perturbations of Dynamical Systems*, Springer, New York.
- [5] Iacus, S. and Kutoyants, Yu. A. (2001) Semiparametric hypotheses testing for dynamical systems with small noise. *Math. Methods Statist.*, 10(1), 105-120.
- [6] Khmaladze, E. (1981) Martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl.*, 26(2), 240-257.
- [7] Kleptsyna, M. and Kutoyants, Yu. A. (2014) On asymptotically distribution free tests with parametric hypothesis for ergodic diffusion processes. *Stat. Inference Stoch. Process.*, 17(3), 295-319.
- [8] Kutoyants, Yu. A. (1978) Estimation of a parameter of a diffusion process. *Theory Probab. Appl.*, 23(3), 641-649.
- [9] Kutoyants, Yu. A. (1994) *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht.
- [10] Kutoyants, Yu. A. (2011) On goodness-of-fit tests for perturbed dynamical systems. *J. Statist. Plann. Inference*, 141(5), 1655-1666.

- [11] Kutoyants, Yu. A. (2014) On ADF goodness-of-fit tests for stochastic processes. *New Perspectives on Stochastic Modeling and Data Analysis*, (J. R. Bozeman, V. Girardin and C. H. Skiadas Eds), ISAST, Athens, 2014, 3-18.
- [12] Kutoyants, Yu. A. (2015) On ADF goodness-of-fit tests for perturbed dynamical systems. *Bernoulli*, 21(4), 2430-2456.
- [13] Kutoyants, Yu. A. (2016) On score-functions and goodness-of-fit tests for stochastic processes. *Math. Methods Statist.*, 25(2), 99-120.
- [14] Lehmann, E. L. and Romano, J. P. (2005) *Testing Statistical Hypotheses*. (3rd ed.) Springer, New York.
- [15] Liptser, R. and Shirayev, A. N. (2005) *Statistics of Random Processes*. v. 2, 2-nd ed. Springer, New York.
- [16] Maglapheridze, N., Tsigroshvili, Z. P. and van Pul, M. (1998) Goodness-of-fit tests tests for parametric hypotheses on the distribution of point processes, *Math. Methods. Statist.* 7, 60-77.
- [17] Negri, I. and Nishiyama, Y. (2009) Goodness of fit test for ergodic diffusion processes. *Ann. Inst. Statist. Math.*, 61(4), 919-928.
- [18] Negri, I. and Zhou, L. (2012) On goodness-of-fit testing for ergodic diffusion process with shift parameter. *Stat. Inference Stoch. Process.*, 17(1), 51-73.